

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 20, 2005, 10:37:17 ; Search time 116 Seconds  
(without alignments)  
1693.744 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPFRHLVLQLALP.....GTRLEIKLYPRGSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqe, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*  
1: geneseqp1980s:\*  
2: geneseqp1990s:\*  
3: geneseqp2000s:\*  
4: geneseqp2001s:\*  
5: geneseqp2002s:\*  
6: geneseqp2003as:\*  
7: geneseqp2003bs:\*  
8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2651.5	99.4	507	3 AAB00158	AAB00158 sCD4-SCFV
2	1064	39.9	295	2 AAR07605	AAR07605 Plasmid T
3	1064	39.9	318	2 AAR07606	AAR07606 Plasmid T
4	1064	39.9	394	2 AAY39825	AAY39825 Soluble h
5	1064	39.9	394	3 AAY88328	AAY88328 T4 glycop
6	1064	39.9	416	3 AAB19509	Aab19509 CD4-IgM f
7	1064	39.9	436	3 AAY51080	AAY51080 Human fus
8	1064	39.9	458	3 AAY88329	Aay88329 T4 glycop
9	1064	39.9	458	4 AAB81502	Aab81502 Human CD4
10	1064	39.9	458	4 AAG79087	Aag79087 Amino aci
11	1064	39.9	458	7 ADD25609	Add25609 Binding d
12	1064	39.9	458	7 ADE57489	Ade57489 Human pro
13	1064	39.9	458	8 ADP12495	Adp12495 Protein e
14	1064	39.9	473	7 ADA44807	Ada44807 CD4/TCR C
15	1064	39.9	474	3 AAY59170	Aay59170 CD4-Ig fu
16	1064	39.9	481	3 AAB19510	Aab19510 CD4-IgM f
17	1064	39.9	481	3 AAY51081	Aay51081 Human fus
18	1064	39.9	481	3 AAY59171	Aay59171 CD4-Ig fu
19	1064	39.9	519	2 AAR20152	AAR20152 Human CD4
20	1064	39.9	616	3 AAY51082	Aay51082 Human fus
21	1064	39.9	616	3 AAY59172	Aay59172 CD4-Ig fu
22	1064	39.9	631	1 AAR93009	Aar93009 Genetic c
23	1064	39.9	631	3 AAB19508	Aab19508 CD4-IgG1
24	1064	39.9	631	3 AAY51079	Aay51079 Human fus
25	1064	39.9	631	3 AAY59169	Aay59169 CD4-Ig fu

26	1064	39.9	729	1 AAP93008	Aap93008 Genetic c
27	1064	39.9	729	3 AAB19507	Aab19507 CD4-IgG1
28	1064	39.9	729	3 AAY51078	Aay51078 Human fus
29	1064	39.9	729	3 AAY59168	Aay59168 CD4-Ig fu
30	1062	39.8	458	7 ADE65841	Ade65841 Human CD4
31	1059	39.7	398	2 AAR78673	Aar78673 CD4 doma
32	1059	39.7	398	2 AAR89450	Aar89450 CD4 DI-D4
33	1059	39.7	400	2 AAR06374	Aar06374 Truncated
34	1059	39.7	402	1 AAP91922	Aap91922 Sequence
35	1059	39.7	402	1 AAP94757	Aap94757 Sequence
36	1059	39.7	458	1 AAP81980	Aap81980 Clone pT4
37	1059	39.7	458	1 AAP91369	Aap91369 T4 protei
38	1059	39.7	458	2 AAR13491	Aar13491 Human CD4
39	1059	39.7	458	2 AAY39826	Aay39826 Soluble h
40	1059	39.7	458	3 AAB07769	Aab07769 DNA encod
41	1059	39.7	462	2 AAR27277	Aar27277 CD4:eta p
42	1059	39.7	462	2 AAR78677	Aar78677 T-cell re
43	1059	39.7	462	2 AAR89457	Aar89457 CD4:gamma
44	1059	39.7	462	2 AAW02214	Aaw02214 CD4:Fc re
45	1059	39.7	462	2 AAW83142	Aaw83142 Chimeric

ALIGNMENTS

RESULT 1	
ID	AAB00158 standard; protein; 507 AA.
XX	
AC	AAB00158;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	sCD4-SCFv(17b) HIV single chain antibody fusion protein.
XX	
KW	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS; -
KW	acquired immune deficiency syndrome; neutralisation; infection;
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;
KW	binding domain; single chain antibody; chimera; chimeric protein.
XX	
OS	Human immunodeficiency virus.
OS	Synthetic.
XX	
PN	MO200055207-A1.
XX	
PD	21-SEP-2000.
XX	
PF	16-MAR-2000; 2000WO-US006946.
XX	
PR	16-MAR-1999; 99US-0124681P.
XX	
PA	(USSH ) US NAT INST OF HEALTH.
XX	
PI	Berger EA, Del Caetlillo CM;
XX	
DR	WPI: 2000-638183/61.
DR	N-PSDB; AAA54045.
XX	
PT	Novel neutralizing bispecific fusion proteins effective in viral such as
PT	HIV neutralization, comprises two different binding domains, inducing-
PT	binding domain and induced-binding domain functionally linked by linker.
XX	
PS	Claim 39; Page 46-47; 55pp; English.
XX	
CC	sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
CC	binding to two sites of its target protein. The protein comprises a first
CC	binding domain capable of binding to an inducing site on the target
CC	protein, a second binding domain capable of forming neutralising complex
CC	with an induced epitope of the target protein and a linker connecting the
CC	binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment
CC	(containing domains D1 and D2) fused to a single chain Fv portion of
CC	antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or
CC	metetic is used for inactivating gp120 protein of HIV, and for

CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure

XX  
 XX Sequence 507 AA;

Query Match 99.4%; Score 2651.5; DB 3; Length 507;  
 Best Local Similarity 99.8%; Pred. No. 9.4e-128;  
 Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MNRGVPFRHLVLTQALLPAATQGGKVVLGKGDVVELTCTASOKKSIQFHKNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAATQGGKVVLGKGDVVELTCTASOKKSIQFHKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQCGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQCGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWTCVTQONKKVEFKIDIVLAFOKASGGGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 240  
 DB 181 TWTCVTQONKKVEFKIDIVLAFOKASGGGSGSGSGSGSGSGSGSGSGSGSGSGSGSG 240  
 QY 241 GGSQVQLLESQAEVKKPSSVKKVSCKASGDTFIRYSFTWROAPGQLEMMGRITITLDV 300  
 DB 241 GGSQVQLLESQAEVKKPSSVKKVSCKASGDTFIRYSFTWROAPGQLEMMGRITITLDV 300  
 QY 301 AHYAPHIQGVTTTADKSTSTVYLELNLRASDDTAVYFCAGVYEGEADDEGEYDNNGLKH 360  
 DB 301 AHYAPHIQGVTTTADKSTSTVYLELNLRASDDTAVYFCAGVYEGEADDEGEYDNNGLKH 360  
 QY 361 WGGGLTVTVSGGGSGSGSGSGSGSGSLELTQSPATLSVSPGERATLSCPASSESVDLA 420  
 DB 361 WGGGLTVTVSGGGSGSGSGSGSGSLELTQSPATLSVSPGERATLSCPASSESVDLA 420  
 QY 421 WYQOKPGQAPRLILYGASTRATGVPARFSGSGSAGEFTLTITISLQSEDPFAVYYCOQYNNV 480  
 DB 421 WYQOKPGQAPRLILYGASTRATGVPARFSGSGSAGEFTLTITISLQSEDPFAVYYCOQYNNV 480  
 QY 481 PPRYTFGGTRLLEIKLVPRGSGHHHHH 508  
 DB 481 PPRYTFGGTRLLEIKLVPRGSGHHHHH 508

RESULT 2  
 AAR07605  
 ID AAR07605 standard; protein; 295 AA.  
 AC AAR07605;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)  
 XX  
 XX Plasmid T4/LFA-3/AD gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 XX Plasma membrane binding affinity; micelle.

XX Unidentified.  
 OS  
 XX Key Location/Qualifiers  
 FH Protein 1..266  
 FT /label= T4/LFA-3/AD gene product  
 FT Peptide 267..295  
 FT /label= PI signal peptide

XX MO9012099-A.  
 XX  
 XX 18-OCT-1990.  
 XX  
 XX 10-APR-1989; 89US-00335688.  
 XX  
 XX 10-APR-1989; 89US-00335688.  
 XX  
 XX (BIOJ) BIOGEN INC.  
 XX  
 XX Wallner BP;  
 XX  
 XX WPI; 1990-334849/44.  
 DR N-PSDB; AAQ06404.  
 PT Phosphatidylinositol linkage signalling DNA sequence - derived from  
 PT lymphocyte function-associated antigen 3, used for prodn. of chimeric  
 PT proteins.  
 XX  
 XX PS Disclosure; Fig 4; 53pp; English.  
 XX  
 XX The signal sequence is attached downstream of the plasmid sequence  
 CC encoding a secreted protein, which will then produce proteins covalently  
 CC anchored to the cell surface in which they were produced. This can give  
 CC rise to plasma membrane binding, enhanced purifiability, micelle  
 CC formation etc. especially useful in the production of chimeric targeted  
 CC drugs; to produce micellar or liposomal delivery systems or in enhanced  
 CC purification and screening of cells, proteins or DNA libraries. (updated  
 CC on 09-JAN-2003 to add missing OS field.) (updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 XX Sequence 295 AA;

Query Match 39.9%; Score 1064; DB 2; Length 295;  
 Best Local Similarity 100.0%; Pred. No. 6.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALLPAATQGGKVVLGKGDVVELTCTASOKKSIQFHKNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPAATQGGKVVLGKGDVVELTCTASOKKSIQFHKNSNOIK 60  
 QY 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 DB 61 ILNGQSFLLTKGSPKLNDRADSRSLMDQGNFLLIKNLKI ESDTYICEVEQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLQCGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQCGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWTCVTQONKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCVTQONKKVEFKIDIVLAFOKAS 208

RESULT 3  
 AAR07606  
 ID AAR07606 standard; protein; 318 AA.  
 AC AAR07606;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)  
 XX

DE Plasmid T4/LFA-3/2 gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 KM Plasma membrane binding affinity; micelle.  
 OS Unidentified.  
 XX  
 FH Key Location/Qualifiers  
 FT Protein 1. 267  
 FT Peptide /label= LFA-3 gene product  
 FT Peptide 267. 318  
 FT /label= PI signal peptide  
 XX  
 PN MO9012099-A.  
 XX  
 PD 18-OCT-1990.  
 XX  
 PF 10-APR-1989; 89US-00335688.  
 XX  
 PR 10-APR-1989; 89US-00335688.  
 XX  
 PA (BIO) BIOGEN INC.  
 XX  
 PI Wallner BP;  
 XX  
 PI WPI; 1990-334849/44.  
 DR N-PSDB; AA060405.  
 XX  
 PT Phosphatidylinositol linkage signalling DNA sequence - derived from  
 PT lymphocyte function-associated antigen 3, used for prodn. of chimeric  
 PT proteins.  
 XX  
 PS Disclosure; Fig 5; 53pp; English.  
 XX  
 CC The signal sequence is attached downstream of the Plasmid sequence  
 CC encoding a secreted protein, which will then produce proteins covalently  
 CC anchored to the cell surface, in which they were produced. This can give  
 CC rise to plasma membrane binding, enhanced purifiability, micelle  
 CC formation etc. especially useful in the production of chimeric targeted  
 CC drugs; to produce micellar or liposomal delivery systems or in enhanced  
 CC purification and screening of cells, proteins or DNA libraries. (Updated  
 CC on 09-JAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to  
 CC correct PA field.)  
 XX  
 SO Sequence 318 AA;  
 XX

Query Match 39.9%; Score 1064; DB 2; Length 318;  
 Best Local Similarity 100.0%; Pred. No. 6.9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVLTQALLPAATQGGKGVLTGKGDVTELTCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLLVLTQALLPAATQGGKGVLTGKGDVTELTCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWTCVTIVONQKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTIVONQKVEFKIDIVLAFQKAS 208

RESULT 4  
 ID AAY39825 standard; procein; 394 AA.  
 XX AAY39825;  
 AC  
 XX

DT 03-DEC-1999 (first entry)  
 XX  
 DE Soluble human T4 protein.  
 XX  
 KM Soluble T4 protein; sT4; human; HIV; binding inhibitor; T4+ cell; AIDS;  
 KM vaccine; immunisation; therapy.  
 XX  
 OS Homo sapiens.  
 XX  
 PN US9598678-A.  
 XX  
 PD 28-SEP-1999.  
 XX  
 PF 12-DEC-1994; 94US-00354452.  
 XX  
 PR 21-AUG-1986; 86US-00898587.  
 PR 11-JUN-1991; 91US-00713564.  
 PR 06-JUL-1992; 92US-00909021.  
 XX  
 PA (UYCO) UNIV COLUMBIA NEW YORK.  
 XX  
 PI McDougal JS, Weiss R, Axel R, Littman DR, Maddon PJ, Chess L;  
 XX  
 PI WPI; 1999-561025/47.  
 DR N-PSDB; AA220694.  
 XX  
 PT Human T4 protein inhibits HIV binding to T4 cells, useful for treating  
 PT AIDS.  
 XX  
 PS Disclosure; Col 13-16; 56pp; English.  
 XX  
 CC This sequence represents the soluble human T4 protein of the invention.  
 CC The soluble human T4 protein blocks the binding of HIV to T4+ cells and  
 CC is therefore useful for the treatment of AIDS. Monoclonal antibodies  
 CC against the T4 protein may be used as vaccines for immunising subjects  
 CC against AIDS  
 XX  
 SO Sequence 394 AA;  
 XX

Query Match 39.9%; Score 1064; DB 2; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVLTQALLPAATQGGKGVLTGKGDVTELTCTASOKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLLVLTQALLPAATQGGKGVLTGKGDVTELTCTASOKKSIOFHMKNNOIK 60  
 QY 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILGNQGSFLTGPSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 QY 181 TWTCVTIVONQKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTIVONQKVEFKIDIVLAFQKAS 208

RESULT 5  
 ID AAY88328 standard; protein; 394 AA.  
 XX AAY88328;  
 AC  
 XX  
 DT 14-JUL-2000 (first entry)  
 XX  
 DE T4 glycoprotein amino acid sequence.  
 XX  
 KM sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
 KM AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
 XX

```

OS Mammalia.
XX US5126433-A.
XX 30-JUN-1992.
XX 23-OCT-1987; 87US-00114244.
XX 21-AUG-1986; 86US-00898587.
XX (UYCO ) UNIV COLUMBIA NEW YORK.
XX
XX Maddon PJ, Chase L, Axel R, Weiss R, Littman DR, McDougal JS;
XX WPI; 2000-348913/30.
XX N-PSDB; AAA10906.
XX
XX Soluble T4 glycoprotein useful for prevention and treatment of acquired
XX immunodeficiency syndrome and for screening inhibitors of human
XX immunodeficiency viral binding.
XX
XX Disclosure; Col 11-16; 64pp; English.
XX
XX This sequence represents the full length amino acid sequence of
XX glycosylated sT4 glycoprotein. Human immunodeficiency virus (HIV) uses
XX sT4 as a target receptor on T cells. The invention relates to
XX glycosylated sT4 which functions by blocking the binding of HIV to T4
XX target cells, and can be used for the prophylaxis and treatment of AIDS
XX patients. Administration of sT4 effectively inhibits the cell to cell
XX spreading of HIV infection and also the fusion of HIV-infected T4 cells
XX and non-infected T4 cells. The administration of T4 alleviates several
XX symptoms associated with AIDS, and prevents the occurrence of new
XX pathological changes. The sT4 glycoprotein is useful for the prophylaxis
XX and treatment of patients with AIDS. It is also useful as a reagent to
XX identify natural, synthetic or recombinant molecules which act as
XX therapeutic agents or inhibitors of T4+ cell interactions and in
XX diagnostic assays for detection T4 proteins or molecules
XX
XX Sequence 394 AA;
XX
XX Query Match 39.9%; Score 1064; DB 3; Length 394;
XX Best Local Similarity 100.0%; Pred. No. 8.3e-47;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MNRGVPFRHLVLYQALPAATQGGKVVLGKGDVVELTCTASQKKSIOFHKNSNOIK 60
XX 1 MNRGVPFRHLVLYQALPAATQGGKVVLGKGDVVELTCTASQKKSIOFHKNSNOIK 60
XX
XX 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEVQL 120
XX 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEVQL 120
XX
XX 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX
XX 181 TWTCTVLQNGKQKVEFKIDIVLAFQKAS 208
XX 181 TWTCTVLQNGKQKVEFKIDIVLAFQKAS 208
XX
XX RESULT 6
XX AAB19509
XX ID AAB19509 standard; protein; 416 AA.
XX
XX AAB19509;
XX
XX 09-JAN-2001 (first entry)
XX
XX CD4-IgM fusion protein CH4Mmu.
XX
XX CD4; IgM, human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;
XX therapy; diagnosis.

```

```

XX OS Homo sapiens.
XX Key Location/Qualifiers
XX FH Protein 1..395 "CD4 extracellular region"
XX FT /note="CD4 extracellular region"
XX FT Protein 400..416
XX FT /note="IgM heavy chain partial sequence"
XX
XX US6117656-A.
XX
XX 12-SEP-2000.
XX
XX 07-JUN-1995; 95US-00479353.
XX
XX 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.
XX PR 04-FEB-1994; 94US-00191708.
XX
XX (GEHO ) GEN HOSPITAL CORP.
XX
XX Seed B;
XX
XX WPI; 2000-586558/55.
XX DR N-PSDB; AAA50662.
XX
XX CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or
XX SIV.
XX
XX Example 1; Col 41-50; 39pp; English.
XX
XX The present sequence is that of fusion protein CD4Mmu comprising the
XX extracellular portion of CD4, which binds to HIV gp120, linked at its C-
XX terminus to the human IgM heavy chain. To obtain the fusion protein, DNA
XX encoding CD4 was linked to IgM DNA at the MscI site upstream of the CH1
XX region (see AAA50662). Fusion protein CD4Mmu and a nucleic acid encoding
XX it are claimed. Also claimed are a vector comprising the nucleic acid,
XX a transformed host cell, the fusion protein in secreted form using a
XX therapeutic agent, radiolabel or NMR imaging agent. The fusion protein
XX can be administered to an animal (including humans) for treatment of HIV
XX or SIV infection, and can also be used in assays for HIV or SIV, imaging
XX and tissue stains. IgM fusion proteins such as CD4Mmu provide complement-
XX mediated immunity
XX
XX Sequence 416 AA;
XX
XX Query Match 39.9%; Score 1064; DB 3; Length 416;
XX Best Local Similarity 100.0%; Pred. No. 8.7e-47;
XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 MNRGVPFRHLVLYQALPAATQGGKVVLGKGDVVELTCTASQKKSIOFHKNSNOIK 60
XX 1 MNRGVPFRHLVLYQALPAATQGGKVVLGKGDVVELTCTASQKKSIOFHKNSNOIK 60
XX
XX 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEVQL 120
XX 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKNLIKIEDSDTYICEVEDQKEVQL 120
XX
XX 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX 121 LVFGLTANSPTHTLLOGQSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
XX
XX 181 TWTCTVLQNGKQKVEFKIDIVLAFQKAS 208
XX 181 TWTCTVLQNGKQKVEFKIDIVLAFQKAS 208
XX
XX RESULT 7
XX AA51080
XX ID AA51080 standard; protein; 436 AA.

```

```

XX AC AAY51080;
XX DT 23-MAR-2000 (first entry)
XX DE Human fusion protein CD4Mg.
XX DE Human fusion protein CD4Mg.
XX DE Fusion protein; human; CD4; Igm; immunoglobulin; gp120;
XX DE anti-human immunodeficiency virus; CD4Mg.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US6004781-A.
XX PD 21-DEC-1999.
XX PF 04-FEB-1994; 94US-00191708.
XX PR 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.
XX PA (GENO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX DR WPI; 2000-085792/07.
XX DR N-PSDB; AAZ44063.
XX PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX PS Example 1; Col 41-50; 39pp; English.
XX CC This invention describes a novel nucleic acid (I) encoding a fusion
XX CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX CC chain (III). The products of the invention have anti-human
XX CC immunodeficiency virus (HIV) activity and are capable of binding to
XX CC gp120. The fusion protein is useful for treating human immunodeficiency
XX CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX CC represents the fusion protein CD4Mg which is constructed from CD4 linked
XX CC to human Igm upstream of the CH1 region
XX SQ Sequence 436 AA;

Query Match 39.9%; Score 1064; DB 3; Length 436;
Best Local Similarity 100.0%; Pred. No. 9.1e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX DT 14-JUL-2000 (first entry)
XX DE T4 glycoprotein amino acid sequence.
XX DE sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;
XX DE AIDS; treatment; inhibit; cell to cell spread; infection; fusion.
XX OS Mammalia.
XX PN US5126433-A.
XX PD 30-JUN-1992.
XX PF 23-OCT-1987; 87US-00114244.
XX PR 21-AUG-1986; 86US-00898587.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;
XX DR WPI; 2000-348913/30.
XX PT Soluble T4 glycoprotein useful for prevention and treatment of acquired
XX PT immunodeficiency syndrome and for screening inhibitors of human
XX PS immunodeficiency viral binding.
XX PS Example; Fig 6; 64pp; English.
XX CC This sequence represents the amino acid sequence of glycosylated sT4
XX CC glycoprotein. Human immunodeficiency virus (HIV) uses sT4 as a target
XX CC receptor on T cells. The invention relates to glycosylated sT4 which
XX CC functions by blocking the binding of HIV to T4 target cells, and can be
XX CC used for the prophylaxis and treatment of AIDS patients. Administration
XX CC of sT4 effectively inhibits the cell to cell spreading of HIV infection
XX CC and also the fusion of HIV-infected T4 cells and non-infected T4 cells.
XX CC The administration of T4 alleviates several symptoms associated with
XX CC AIDS, and prevents the occurrence of new pathological changes. The sT4
XX CC glycoprotein is useful for the prophylaxis and treatment of patients with
XX CC AIDS. It is also useful as a reagent to identify natural, synthetic or
XX CC recombinant molecules which act as therapeutic agents or inhibitors of
XX CC T+ cell interactions and in diagnostic assays for detection T4 proteins
XX CC or molecules
XX SQ Sequence 458 AA;

Query Match 39.9%; Score 1064; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 9.5e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

RESULT 8
AAY88329
ID AAY88329 standard; protein; 458 AA.
XX
AC AAY88329;

```

```

RESULT 9
AAB81502
ID AAB81502 standard; protein; 458 AA.
XX
AC AAB81502;

```

XX 18-JUN-2001 (first entry)  
 XX Human CD4 protein.  
 DE Human CD4 protein.  
 XX Human; CD4; CD4 fusion protein; oligomerisation;  
 KM receptor-ligand interaction inhibition; surface plasmon resonance; SPR;  
 KM T cell receptor binding; MHC binding; carcinoma; autoimmune disease;  
 KM multiple sclerosis; human immunodeficiency virus; HIV; diabetes;  
 KM rheumatoid arthritis; immune disorder.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 FH Key  
 FT Peptide 1..25  
 FT /label= Signal\_peptide  
 FT Protein 26..458  
 FT /label= Human\_CD4  
 FT MO200122084-A2.  
 PN 29-MAR-2001.  
 PD 29-MAR-2001.  
 PF 18-SEP-2000; 2000WO-GB003579.  
 XX 21-SEP-1999; 99GB-00022352.  
 PR 21-SEP-1999; 99GB-00022352.  
 XX (AVID-) AVIDEX LTD.  
 PA (AVID-) AVIDEX LTD.  
 XX Jakobson BK;  
 PI Jakobson BK;  
 DR WPI; 2001-273470/28.  
 XX N-PSDB; AAF82582.  
 XX Sequential screening of candidate compound library for those which  
 PT inhibit binding of low affinity receptor-ligand interaction having fast  
 PT binding kinetics, using interfacial optical assay.  
 XX Disclosure; Fig 13; 91pp; English.  
 PS The present sequence is human CD4. Human CD4 extracellular domains 1 and  
 XX 2 were used in the construction of CD4 oligomerisation fusion proteins.  
 CC The fusion proteins contain an oligomerisation domain that enables the  
 CC proteins to bind to one another to form oligomers. The oligomers may be  
 CC used in an invention relating to a method for screening for compounds  
 CC with the ability to inhibit a low affinity receptor-ligand interaction.  
 CC The method uses an interfacial optical assay, such as surface plasmon  
 CC resonance (SPR). The method is useful for screening candidate compounds  
 CC for the ability to inhibit interaction between MHC/peptide complex and T  
 CC cell receptor, and MHC/peptide complex and CD8 or CD4 co-receptor. The  
 CC compounds identified by the above methods which interfere with T cell  
 CC receptor binding to a particular HLA type molecule are useful as immune  
 CC inhibitors for treating carcinomas, autoimmune diseases such as multiple  
 CC sclerosis, human immunodeficiency virus (HIV) infection, rheumatoid  
 CC arthritis, Hashimoto's disease, insulin dependent diabetes, Good  
 CC pasture's syndrome, uveitis, psoriasis and graft rejection  
 CC XX  
 SQ Sequence 458 AA;  
 Query Match 39.9%; Score 1064; DB 4; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
 DB 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
 QY 1 ILNQGSGFLTKGSKLNDRADSRSLWDQGNFPLIIRNLIKIEBDDTYICEVEDQKEEVL 120  
 DB 61 ILNQGSGFLTKGSKLNDRADSRSLWDQGNFPLIIRNLIKIEBDDTYICEVEDQKEEVL 120  
 QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCSPRGKNIQGKTLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCSPRGKNIQGKTLVSQLELDGSG 180

DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCSPRGKNIQGKTLVSQLELDGSG 180  
 QY 181 TWCTCTVLQONQKVEFKIDIVLAFOKAS 208  
 DB 181 TWCTCTVLQONQKVEFKIDIVLAFOKAS 208  
 RESULT 10  
 ID AAG79087  
 ID AAG79087 standard; protein; 458 AA.  
 XX AAG79087;  
 AC AAG79087;  
 XX 10-DEC-2001 (first entry)  
 DT 10-DEC-2001 (first entry)  
 DE Amino acid sequence of human CD4 protein.  
 XX Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;  
 KM C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; CD4.  
 XX Homo sapiens.  
 OS Homo sapiens.  
 PN MO200164752-A2.  
 PD 07-SEP-2001.  
 PF 28-FEB-2001; 2001WO-US006322.  
 XX 02-MAR-2000; 2000US-00517605.  
 PR (UYNV ) UNIV NEW YORK STATE.  
 PA (UYNV-) UNIV NIJMEGEN.  
 XX Liltman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;  
 PI Liltman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;  
 DR WPI; 2001-602565/68.  
 XX An antibody for the treatment or prevention of HIV-infection comprises a  
 PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
 PT DC-SIGN due to concomitant conformational change.  
 XX Disclosure; Page 115-116; 131pp; English.  
 PS The specification describes an antibody which is specific for an  
 XX antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
 CC is exposed upon gp120 binding of DC-SIGN due to concomitant  
 CC conformational change. DC-SIGN is a receptor that is specifically  
 CC expressed on dendritic cells and facilitates infection of T lymphocytes  
 CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
 CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
 CC with high affinity. The antibody of the invention inhibits the trans  
 CC enhancement of HIV entry into a T cell or macrophage facilitated by  
 CC dendritic cells. The antibody is useful to treat or prevent HIV  
 CC infection. The present sequence represents a human CD4 protein  
 CC XX  
 SQ Sequence 458 AA;  
 Query Match 39.9%; Score 1064; DB 4; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 9.5e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
 DB 1 MNRGVPFRHLILVQLALLPAATGKKVVLGKKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
 QY 1 ILNQGSGFLTKGSKLNDRADSRSLWDQGNFPLIIRNLIKIEBDDTYICEVEDQKEEVL 120  
 DB 61 ILNQGSGFLTKGSKLNDRADSRSLWDQGNFPLIIRNLIKIEBDDTYICEVEDQKEEVL 120  
 QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCSPRGKNIQGKTLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCSPRGKNIQGKTLVSQLELDGSG 180

QY 181 TWCTVLQNGKVEFKIDIVLAFOKAS 208  
 DB 181 TWCTVLQNGKVEFKIDIVLAFOKAS 208

RESULT 11  
 ADD25609

ID ADD25609 standard; protein; 458 AA.

XX ADD25609;

XX 15-JAN-2004 (first entry)

DE Binding domain-immunoglobulin fusion protein-associated protein #82.

XX Binding domain; immunoglobulin; fusion protein; cytostatic;  
 KW antiarthritic; immunosuppressive; antidiabetic; antithyroid;  
 KW neuroprotective; hinge region; immunoglobulin heavy chain;  
 KW CH2 constant region; CH3 constant region; IgG1;  
 KW antibody dependent cell-mediated cytotoxicity; ADCC; complement fixation;  
 KW malignant condition; B-cell disorder; melanoma; carcinoma; sarcoma;  
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
 KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.  
 OS Unidentified.

XX US2003118592-A1.

XX 26-JUN-2003.

PF 25-JUL-2002; 2002US-00207655.

PR 17-JAN-2001; 2001US-0367358P.

PR 17-JUN-2002; 2002US-00053530.

PR 03-JUN-2002; 2002US-0385691P.

(GENE-) GENE-CRAFT INC.

PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;

DR WPI, 2003-801317/75.

PT New binding domain-immunoglobulin fusion protein, useful for treating a  
 PT subject having or suspected of having a malignant condition or a B-cell  
 PT disorder, e.g. melanoma, Grave's disease or autoimmune disease.

PS Disclosure; SEQ ID NO 170; 157pp; English.

XX The invention relates to a binding domain-immunoglobulin fusion protein  
 CC comprising a binding domain polypeptide that is fused to an  
 CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
 CC CH2 constant region polypeptide that is fused to the hinge region  
 CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
 CC polypeptide that is fused to the CH2 constant region polypeptide. The  
 CC hinge region polypeptide comprises a wild-type human IgG1 immunoglobulin  
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
 CC region polypeptide; derived from (a) having 3 or more cysteine residues;  
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
 CC mutated human IgG1 immunoglobulin hinge region polypeptide; derived from  
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
 CC immunoglobulin hinge region polypeptide contains no more than one  
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
 CC polypeptide; derived from (a) having 3 or more cysteine residues; where  
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
 CC capable of at least one immunological activity comprising antibody  
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
 CC binding domain polypeptide is capable of specifically binding to an  
 CC antigen. Also included are an isolated polynucleotide encoding the  
 CC binding domain-immunoglobulin fusion protein, a recombinant expression  
 CC construct comprising the polynucleotide (operably linked to a promoter),  
 CC a host cell transformed or transfected with a recombinant expression

CC construct, producing the binding domain-immunoglobulin fusion protein, a  
 CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g. melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type 1 diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC seqdata.uspto.gov/sequence.html?docID=20030118592. The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.

XX Sequence 458 AA;

XX Query Match 39.9%; Score 1064; DB 7; Length 458;

XX Best Local Similarity 100.0%; Pred. No. 9, 5e-47;

XX Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPPRHLLVQLALPAATQGGKVVLGKGGTVELCTRASQKKSIOFHKNNSQIK 60

DB 1 MNRGVPPRHLLVQLALPAATQGGKVVLGKGGTVELCTRASQKKSIOFHKNNSQIK 60

QY 61 ILNGQSFLLTKGSKNDRAISRSLMDQGNPFLIKNLKIEPSDYICEVEPOKREYOL 120

DB 61 ILNGQSFLLTKGSKNDRAISRSLMDQGNPFLIKNLKIEPSDYICEVEPOKREYOL 120

QY 121 LVFGLTRANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

DB 121 LVFGLTRANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180

QY 181 TWCTVLQNGKVEFKIDIVLAFOKAS 208

DB 181 TWCTVLQNGKVEFKIDIVLAFOKAS 208

RESULT 12  
 ADE57489

XX ID ADE57489 standard; protein; 458 AA.

XX ADE57489;

XX 29-JAN-2004 (first entry)

DE Human Protein P01730, SEQ ID NO 3351.

KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SN1; Chung.

XX Homo sapiens.

PN WO2003016475-A2.

PD 27-FEB-2003.

PF 14-AUG-2002; 2002WO-US025765.

PR 14-AUG-2001; 2001US-0312147P.

PR 01-NOV-2001; 2001US-0346382P.

PR 26-NOV-2001; 2001US-033347P.

PA (GENO) GEN HOSPITAL CORP.

PA (PARB) BAYER AG.

PI Woolf C, D'urso D, Befort K, Costigan M;

DR WPI, 2003-268312/26.

DR GENBANK; P01730.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.  
XX  
PS Claim 1; Page; 1017p; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g. spinal segmental nerve injury (SNII), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g. gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.  
XX

SQ Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 7; Length 458;

Best Local Similarity 100.0%; Pred. No. 9.5e-47;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNRGVPRHLLVQLALPAATQGGKVVLGKGGDTVELCTASQKKSIOFHKNSNQIK	60
DB	1	MNRGVPRHLLVQLALPAATQGGKVVLGKGGDTVELCTASQKKSIOFHKNSNQIK	60
QY	61	ILNGQSFLLTKGPKSKNDRAISRSLMDQGNFLLIKNLIKEDSDTYICEVEDQKEEVQL	120
DB	61	ILNGQSFLLTKGPKSKNDRAISRSLMDQGNFLLIKNLIKEDSDTYICEVEDQKEEVQL	120
QY	121	LVFGLTANSTHLLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG	180
DB	121	LVFGLTANSTHLLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG	180
QY	181	TWTCVTVLQNGKVEFKIDIVLAFOKAS	208
DB	181	TWTCVTVLQNGKVEFKIDIVLAFOKAS	208

RESULT 13

ID ADP12495 standard; protein; 458 AA.

AC ADP12495;

DT 12-AUG-2004 (first entry)

DE Protein encoded by mRNA of the invention #105.

XX transplant rejection; immune system; rheumatoid arthritis; lupus;

KW inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

XX Homo sapiens.

XX WO2004042346-A2.

XX

PD 21-MAY-2004.

XX 24-APR-2003; 2003WO-US012946.

XX 24-APR-2002; 2002US-00131831.

XX 20-DEC-2002; 2002US-00325899.

XX (EXPR-) EXPRESSION DIAGNOSTICS INC.

XX Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;

PI Rosenberg S;

XX WPI; 2004-400724/37.

PT Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.

PS Claim 65; SEQ ID NO 2504; 1762p; English.

XX The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprising detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The method is also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.

SQ Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 8; Length 458;

Best Local Similarity 100.0%; Pred. No. 9.5e-47;

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNRGVPRHLLVQLALPAATQGGKVVLGKGGDTVELCTASQKKSIOFHKNSNQIK	60
DB	1	MNRGVPRHLLVQLALPAATQGGKVVLGKGGDTVELCTASQKKSIOFHKNSNQIK	60
QY	61	ILNGQSFLLTKGPKSKNDRAISRSLMDQGNFLLIKNLIKEDSDTYICEVEDQKEEVQL	120
DB	61	ILNGQSFLLTKGPKSKNDRAISRSLMDQGNFLLIKNLIKEDSDTYICEVEDQKEEVQL	120
QY	121	LVFGLTANSTHLLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG	180
DB	121	LVFGLTANSTHLLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG	180
QY	181	TWTCVTVLQNGKVEFKIDIVLAFOKAS	208
DB	181	TWTCVTVLQNGKVEFKIDIVLAFOKAS	208

RESULT 14

ID ADA44807 standard; protein; 473 AA.

AC ADA44807;

DT 04-DEC-2003 (first entry)

DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon15, SEQ ID NO:2.

XX HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;

KW endoplasmic reticulum; ER retention; envelope protein gp160;

KW T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon15;

XX gene therapy; human; receptor.

XX

OS Chimeric.  
 XX Homo sapiens.  
 FH Key Location/Qualifiers  
 FT Protein 1..458  
 FT /label= CD4  
 FT Region 459..473  
 FT /note= "Part of the C-terminal domain of the T cell  
 FT receptor CD3epsilon chain"  
 XX  
 XX W02003076468-A1.  
 XX  
 XX 18-SEP-2003.  
 XX  
 XX 14-MAR-2003; 2003WO-ES000120.  
 XX  
 XX 14-MAR-2002; 2002ES-00000616.  
 XX  
 XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 XX  
 XX Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
 PI Gomez Buendia M;  
 DR WPI; 2003-779059/73.  
 DR N-PSDB; ADA44806.  
 XX  
 XX Composition for treating or preventing human immune deficiency virus,  
 PT comprises CD4 chimeric protein having a protective effect in trans, or  
 PT related nucleic acid.  
 XX  
 XX Claim 5; Page 33-35; 43pp; Spanish.  
 XX  
 CC The invention relates to a composition for the treatment or prevention of  
 CC human immunodeficiency virus-1 (HIV-1) infection. The composition  
 CC comprises CD4+ cells that have been transduced with a vector that encodes  
 CC a chimeric CD4 molecule which is capable of being retained in the  
 CC endoplasmic reticulum (ER). The invention also encompasses the use of a  
 CC soluble protein factor produced by CD4+ cells that have been transduced  
 CC with a vector encoding a chimeric CD4 protein; and the use of an  
 CC expression system encoding a chimeric CD4 protein. The ER-localised  
 CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
 CC resulting in HIV-1 retention in the ER and thereby preventing viral  
 CC replication. In a specific embodiment, the chimeric CD4 molecule  
 CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
 CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
 CC CD4epsilon15 (ADA44807). A known chimeric CD4 of similar structure but  
 CC containing only 10 amino acids from CD3epsilon can also be used.  
 CC Compositions of the invention have an in trans effect on the replication  
 CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
 CC present sequence represents the chimeric CD4 molecule CD4epsilon15, which  
 CC is specifically claimed for use in compositions of the invention.  
 XX  
 XX Sequence 473 AA;  
 SQ  
 Query Match 39.9%; Score 1064; DB 7; Length 473;  
 Best Local Similarity 100.0%; Pred. No. 9, 8e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15  
 ID AAY59170  
 AC AAY59170;  
 XX  
 XX 14-MAR-2000 (first entry)  
 DT  
 XX  
 DE CD4-Ig fusion protein CD4Mmu.  
 XX  
 XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 XX secreted protein; SIV infection; medicament.  
 KM  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX CA1340741-C.  
 PN  
 XX 14-SEP-1999.  
 PD  
 XX 20-JAN-1989; 89CA-00588749.  
 PF  
 XX 20-JAN-1989; 89CA-00588749.  
 PR  
 XX 20-JAN-1989; 89CA-00588749.  
 XX  
 PA (GENO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 PT  
 XX WPI; 2000-063015/06.  
 DR N-PSDB; AA248203.  
 DR  
 XX  
 PT New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 XX Example 1; Page 47-53; 89pp; English.  
 XX  
 CC The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the fusion protein  
 CC CD4Mmu where the CD4 is linked to human IgG1 at the Met2 site upstream of  
 CC the CH1 region  
 XX  
 XX Sequence 474 AA;  
 SQ  
 Query Match 39.9%; Score 1064; DB 3; Length 474;  
 Best Local Similarity 100.0%; Pred. No. 9, 8e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Search completed: September 20, 2005, 16:07:08  
Job time : 121 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

## OM protein - protein search, using SW model

Run on: September 20, 2005, 15:24:08 ; Search time 32 Seconds  
(without alignments)  
1185.054 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPFRHLVLVQLALP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/1/iaa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	DB ID	Description
1	1064	39.9	295	6 5223394-9	Patent No. 5223394
2	1064	39.9	295	6 5223394-9	Patent No. 5223394
3	1064	39.9	318	6 5223394-11	Patent No. 5223394
4	1064	39.9	318	6 5223394-11	Patent No. 5223394
5	1064	39.9	458	3 US-09-039-555B-15	Sequence 15, Appl
6	1064	39.9	458	3 US-09-517-605-3	Sequence 3, Appl
7	1062	39.8	458	4 US-10-092-138A-25	Sequence 25, Appl
8	1059	39.7	398	2 US-08-284-391B-29	Sequence 29, Appl
9	1059	39.7	398	4 US-09-218-950-29	Sequence 29, Appl
10	1059	39.7	398	4 US-08-394-388A-29	Sequence 29, Appl
11	1059	39.7	402	1 US-08-236-311-1	Sequence 1, Appl
12	1059	39.7	402	1 US-08-457-918-1	Sequence 1, Appl
13	1059	39.7	402	4 US-10-157-408-1	Sequence 1, Appl
14	1059	39.7	457	4 US-08-328-500-9	Sequence 9, Appl
15	1059	39.7	458	3 US-08-466-368-4	Sequence 4, Appl
16	1059	39.7	462	2 US-08-417-495-5	Sequence 5, Appl
17	1059	39.7	462	2 US-08-284-391B-5	Sequence 5, Appl
18	1059	39.7	462	3 US-09-218-950-5	Sequence 5, Appl
19	1059	39.7	462	4 US-08-394-388A-5	Sequence 5, Appl
20	1059	39.7	462	4 PCT-US92-01785-5	Sequence 5, Appl
21	1059	39.7	462	5 PCT-US95-00454-5	Sequence 5, Appl
22	1059	39.7	532	2 US-08-417-495-6	Sequence 6, Appl
23	1059	39.7	532	2 US-08-284-391B-6	Sequence 6, Appl
24	1059	39.7	532	2 US-09-218-950-6	Sequence 6, Appl
25	1059	39.7	532	4 US-08-394-388A-6	Sequence 6, Appl
26	1059	39.7	532	4 PCT-US92-01785-6	Sequence 6, Appl
27	1059	39.7	532	5 PCT-US95-00454-6	Sequence 6, Appl

28	1059	39.7	575	2 US-08-417-495-4	Sequence 4, Appl
29	1059	39.7	575	2 US-08-284-391B-4	Sequence 4, Appl
30	1059	39.7	575	3 US-09-218-950-4	Sequence 4, Appl
31	1059	39.7	575	4 US-08-394-388A-4	Sequence 4, Appl
32	1059	39.7	575	5 PCT-US92-01785-4	Sequence 4, Appl
33	1059	39.7	575	5 PCT-US95-00454-4	Sequence 4, Appl
34	1059	39.7	630	4 US-08-472-888A-6	Sequence 6, Appl
35	1056	39.6	458	6 5223394-7	Patent No. 5223394
36	1056	39.6	458	6 5223394-7	Patent No. 5223394
37	1053	39.5	394	4 US-08-466-368-2	Sequence 2, Appl
38	1053	39.5	394	4 US-08-328-500-2	Sequence 2, Appl
39	1051	39.4	530	3 US-08-477-460B-4	Sequence 4, Appl
40	1051	39.4	530	3 US-08-379-516-4	Sequence 4, Appl
41	1051	39.4	530	3 US-09-329-916-4	Sequence 4, Appl
42	1051	39.4	530	3 US-08-485-372A-4	Sequence 4, Appl
43	1051	39.4	530	3 US-09-409-006A-4	Sequence 4, Appl
44	1051	39.4	530	4 US-08-484-681-4	Sequence 4, Appl
45	1051	39.4	530	4 US-09-766-995-4	Sequence 4, Appl

## ALIGNMENTS

RESULT 1  
5223394-9  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO:9:  
; LENGTH: 295  
5223394-9

Query Match 39.9%; Score 1064; DB 6; Length 295;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPAAATQKKVVLGGKGDVETLTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPFRHLVLVQLALPAAATQKKVVLGGKGDVETLTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEQKEVOL 120  
DB 61 ILGNQGSFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEQKEVOL 120  
QY 121 LVGGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPGKNTQGGKTLISVSLLEIDSG 180  
DB 121 LVGGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPGKNTQGGKTLISVSLLEIDSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
RESULT 2  
5223394-9  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO:9:  
; LENGTH: 295  
5223394-9

Query Match 39.9%; Score 1064; DB 6; Length 295;  
Best Local Similarity 100.0%; Pred. No. 9.1e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208

RESULT 3  
5223394-11

APPLICANT: WALLNER, BARBARA  
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
LVMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
LINKAGE SIGNAL SEQUENCE  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/335,688  
FILING DATE: 10-APR-1989  
SEQ ID NO:11:  
LENGTH: 318

5223394-11

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9.8e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208

RESULT 4  
5223394-11

APPLICANT: WALLNER, BARBARA  
TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
LVMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
LINKAGE SIGNAL SEQUENCE  
NUMBER OF SEQUENCES: 12  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/07/335,688  
FILING DATE: 10-APR-1989  
SEQ ID NO:11:  
LENGTH: 318

5223394-11

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9.8e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
DB 61 ILGNQGSFLTKGFSKLNDRADSRSLMDQGNFLLIKNLKIBSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208  
DB 181 TWTCTVLQONOKVFEKIDIVLAFQKAS 208

RESULT 5

US-09-039-555B-15  
Sequence 15, Application US/09039555B

PATENT No. 6033856  
GENERAL INFORMATION:  
APPLICANT: Koerner, Kathrin  
APPLICANT: Mueller, Rolf  
TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
NUMBER OF SEQUENCES: 19  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Foley & Lardner  
STREET: 3000 K Street, N.W., Suite 500  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20007-5109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/039,555B  
FILING DATE: 16-MAR-1998  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: DE 19710643.9  
FILING DATE: 14-MAR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Bent, Stephen A.  
REGISTRATION NUMBER: 29,768  
REFERENCE/DOCKET NUMBER: 016779/0131  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202)672-5300  
TELEFAX: (202)672-5399  
TELEX: 904136  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-039-555B-15

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.4e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATGGKVVLGKKDVTVELTCTASQKKSIOFHWKNSNOIK 60

```
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKGDVTELCTASQKKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208
Db 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 6
US-09-517-605-3
; Sequence 3, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Liteman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 3
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-517-605-3

Query Match 39.9%; Score 1064; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 1,4e-53;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKGDVTELCTASQKKSIOFHWKNSNOIK 60
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKGDVTELCTASQKKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208
Db 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 7
US-10-092-138A-25
; Sequence 25, Application US/10092138A
; Patent No. 6743630
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138A
; CURRENT FILING DATE: 2002-03-06
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 25
```

```
; LENGTH: 458
; TYPE: PRT
; ORGANISM: human
US-10-092-138A-25

Query Match 39.8%; Score 1062; DB 4; Length 458;
Best Local Similarity 99.5%; Pred. No. 1.8e-53;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKGDVTELCTASQKKSIOFHWKNSNOIK 60
Db 1 MNRGVPFRHLLVQLALLPAATQGGKVVLGKGDVTELCTASQKKSIOFHWKNSNOIK 60
QY 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
Db 61 ILGNQGSFLTKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db 121 LVFGLTANSDTHLLQGSGLTITLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208
Db 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 8
US-08-284-391B-29
; Sequence 29, Application US/08284391B
; Patent No. 5851828
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391B
; FILING DATE: 02-AUG-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L.
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
; TELEX:
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
```

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-29

Query Match 39.7%; Score 1059; DB 2; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2,3e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATQKRVLGKGDVVELCTASQKSIQFHWKNSNQIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQKRVLGKGDVVELCTASQKSIQFHWKNSNQIK 60  
QY 61 ILNQGSLFKTGPSKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVL 120  
DB 61 ILNQGSLFKTGPSKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTANSDTHLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCTVLQNKVKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKVKVEFKIDIVLAFQKAS 208

## RESULT 9

US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6284240

## GENERAL INFORMATION:

APPLICANT: Seed, Brian  
APPLICANT: Banepour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:

## CLASSIFICATION:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001

## TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:

LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

Query Match 39.7%; Score 1059; DB 3; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2,3e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAATQKRVLGKGDVVELCTASQKSIQFHWKNSNQIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQKRVLGKGDVVELCTASQKSIQFHWKNSNQIK 60  
QY 61 ILNQGSLFKTGPSKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVL 120  
DB 61 ILNQGSLFKTGPSKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEVL 120  
QY 121 LVFGLTANSDTHLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLQGSLLTLLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCTVLQNKVKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKVKVEFKIDIVLAFQKAS 208

## RESULT 10

US-08-394-388A-29  
Sequence 29, Application US/08394388A  
Patent No. 6753162

## GENERAL INFORMATION:

APPLICANT: Seed, Brian  
APPLICANT: Banepour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995

## CLASSIFICATION:

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L

REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:

INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-388A-29

Query Match 39.7%; Score 1059; DB 4; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.3e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPLPATQKKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPFRHLVLLVQLALPLPATQKKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLONQKVEFKIDIVVLAFOKAS 208  
DB 181 TWICTVLONQKVEFKIDIVVLAFOKAS 208

## RESULT 11

US-08-236-311-1

Sequence 1, Application US/08236311

Patent No. 5565335

GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.

APPLICANT: Gregory, Timothy J.

TITLE OF INVENTION: Adhesion Variants

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 KB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/236.311

FILING DATE: 02-MAY-1994

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/936190

FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/842777

FILING DATE: 18-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/250785

FILING DATE: 28-SEP-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/104329

FILING DATE: 02-OCT-1987

ATTORNEY/AGENT INFORMATION:

NAME: Haasek, Janet E.

REGISTRATION NUMBER: 28,616

REFERENCE/DOCKET NUMBER: 444PIC2

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1896

TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-1

Query Match 39.7%; Score 1059; DB 1; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.4e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLLVQLALPLPATQKKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVPFRHLVLLVQLALPLPATQKKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLONQKVEFKIDIVVLAFOKAS 208  
DB 181 TWICTVLONQKVEFKIDIVVLAFOKAS 208

## RESULT 12

US-08-457-918-1

Sequence 1, Application US/08457918

Patent No. 617655

GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.

APPLICANT: Gregory, Timothy J.

TITLE OF INVENTION: Adhesion Variants

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 KB floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/457.918

FILING DATE: 1-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/236311

FILING DATE: 02-MAY-1994

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/936190

FILING DATE: 26-AUG-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/842777

FILING DATE: 18-FEB-1992

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/250785

FILING DATE: 28-SEP-1988

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/104329

FILING DATE: 02-OCT-1987

ATTORNEY/AGENT INFORMATION:

NAME: Kubienc, Jeffrey S.

REGISTRATION NUMBER: 36,575

REFERENCE/DOCKET NUMBER: P0444PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-1

Query Match 39.7%; Score 1059; DB 3; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2,4e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGRKVVLGKGDVLELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATQGRKVVLGKGDVLELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTANSDTHLLQGSSTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSSTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 13  
US-10-157-408-1

Sequence 1, Application US/10157408  
Patent No. 6710169

GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.

Gregory, Timothy J.

TITLE OF INVENTION: Adhesion Variants

NUMBER OF SEQUENCES: 25

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genentech, Inc.

STREET: 460 Point San Bruno Blvd

CITY: South San Francisco

STATE: California

COUNTRY: USA

ZIP: 94080

COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: patin (Genentech)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/157,408

FILING DATE: 28-May-2002

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/457,918

FILING DATE: 1-JUN-1995

APPLICATION NUMBER: 08/236311

FILING DATE: 02-MAY-1994

APPLICATION NUMBER: 07/936190

FILING DATE: 26-AUG-1992

APPLICATION NUMBER: 07/842777

FILING DATE: 18-FEB-1992

APPLICATION NUMBER: 07/250785

FILING DATE: 28-SEP-1988

APPLICATION NUMBER: 07/104329

FILING DATE: 02-OCT-1987

ATTORNEY/AGENT INFORMATION:

NAME: Kubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444PIC3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-10-157-408-1

Query Match 39.7%; Score 1059; DB 4; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2,4e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGRKVVLGKGDVLELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATQGRKVVLGKGDVLELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILNGQSFLLTKGSKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTANSDTHLLQGSSTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSSTLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 14  
US-08-328-500-9

Sequence 9, Application US/08328500

Patent No. 6673896

GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.

APPICANT: Axel, Richard

APPICANT: Sweet, Richard W.

APPICANT: Archos, James

TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4

NUMBER OF SEQUENCES: 22

CORRESPONDENCE ADDRESS:

ADDRESSEE: Cooper & Dunham

STREET: 1185 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: U.S.A.

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/328,500

FILING DATE:

CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

NAME: White, John P.

REGISTRATION NUMBER: 28,678

REFERENCE/DOCKET NUMBER: 0575/24577-CY

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 278-0400

TELEFAX: (212) 391-0525

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 457 amino acids

TYPE: amino acid

TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-9

Query Match 39.7%; Score 1059; DB 4; Length 457;  
Best Local Similarity 99.5%; Pred. No. 2.7e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVFRRHLLVQLALPPATQGGKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVFRRHLLVQLALPPATQGGKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208

## RESULT 15

US-08-466-368-4  
Sequence 4, Application US/08466368  
Patent No. 6093539

## GENERAL INFORMATION:

APPLICANT: Maddon, Paul J.  
APPLICANT: Littman, Dan R.  
APPLICANT: Chess, Leonard  
APPLICANT: Axel, Richard  
APPLICANT: Weiss, Robin  
APPLICANT: McDougal, J. S.  
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036

## COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,368  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435

## ATTORNEY/AGENT INFORMATION:

NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 24577-E1-B/JPM/AKC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-278-0400  
TELEFAX: 212-391-0525  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-466-368-4

Query Match 39.7%; Score 1059; DB 3; Length 458;  
Best Local Similarity 99.5%; Pred. No. 2.7e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVFRRHLLVQLALPPATQGGKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
DB 1 MNRGVFRRHLLVQLALPPATQGGKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60  
QY 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPKSKINDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208  
DB 181 TWICTVLQNGKVEFKIDIVLAFQKAS 208

Search completed: September 20, 2005, 16:10:48  
Job time: 34 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 16:09:38 ; Search time 485 Seconds  
(without alignments)  
424.130 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPFRHLLVLQALALP.....GTRLEIKLVPRSGSHHHHH 508

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1812044 seqs, 404927589 residues

Total number of hits satisfying chosen parameters: 1812044

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: /cgnt\_6/prodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgnt\_6/prodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgnt\_6/prodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
4: /cgnt\_6/prodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
5: /cgnt\_6/prodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
6: /cgnt\_6/prodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgnt\_6/prodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgnt\_6/prodata/1/pubppaa/US08\_PUBCOMB.pep:\*  
9: /cgnt\_6/prodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
10: /cgnt\_6/prodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
11: /cgnt\_6/prodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgnt\_6/prodata/1/pubppaa/US09C\_NEW\_PUB.pep:\*  
13: /cgnt\_6/prodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgnt\_6/prodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
15: /cgnt\_6/prodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgnt\_6/prodata/1/pubppaa/US10D\_PUBCOMB.pep:\*  
17: /cgnt\_6/prodata/1/pubppaa/US10E\_PUBCOMB.pep:\*  
18: /cgnt\_6/prodata/1/pubppaa/US10E\_NEW\_PUB.pep:\*  
19: /cgnt\_6/prodata/1/pubppaa/US11\_PUBCOMB.pep:\*  
20: /cgnt\_6/prodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
21: /cgnt\_6/prodata/1/pubppaa/US60\_PUBCOMB.pep:\*  
22: /cgnt\_6/prodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1064	39.9	458 14	US-10-151-274-3
2	1064	39.9	458 14	US-10-103-597A-39
3	1064	39.9	458 14	US-10-188-444-39
4	1064	39.9	458 14	US-10-207-655-170
5	1062	39.8	458 8	US-08-681-219-27
6	1062	39.8	458 10	US-09-230-111C-25
7	1062	39.8	458 14	US-10-092-138-25
8	1062	39.8	458 16	US-10-820-403-25
9	1059	39.7	458 16	US-10-939-537-29
10	1059	39.7	402 14	US-10-157-408-1
11	1059	39.7	402 14	US-10-097-044A-1

12	1059	39.7	402 16	US-10-769-247-1	Sequence 1, App11
13	1059	39.7	457 11	US-09-891-119A-9	Sequence 9, App11
14	1059	39.7	462 10	US-09-939-537-5	Sequence 5, App11
15	1059	39.7	462 11	US-09-243-008-5	Sequence 5, App11
16	1059	39.7	532 10	US-09-939-537-6	Sequence 6, App11
17	1059	39.7	532 11	US-09-243-008-6	Sequence 6, App11
18	1059	39.7	575 10	US-09-939-537-4	Sequence 4, App11
19	1059	39.7	575 11	US-09-243-008-4	Sequence 4, App11
20	1053	39.5	397 8	US-09-891-119A-2	Sequence 2, App11
21	1051	39.4	530 11	US-08-485-163-5	Sequence 5, App11
22	1051	39.4	530 9	US-09-766-995-4	Sequence 4, App11
23	1050	39.4	432 8	US-08-485-163-3	Sequence 3, App11
24	1050	39.4	432 9	US-09-766-995-2	Sequence 2, App11
25	1048	39.3	310 8	US-08-485-163-7	Sequence 7, App11
26	1048	39.3	310 9	US-09-766-995-6	Sequence 6, App11
27	1040	39.0	203 18	US-10-513-725-10	Sequence 10, App1
28	1040	39.0	203 16	US-10-512-966-7	Sequence 7, App11
29	1040	39.0	449 16	US-10-493-676-11	Sequence 11, App1
30	1040	39.0	450 16	US-10-493-676-6	Sequence 6, App11
31	1040	39.0	487 18	US-10-512-966-8	Sequence 8, App11
32	1036	38.8	202 16	US-10-492-729-19	Sequence 19, App1
33	1035	38.8	203 10	US-09-939-537-31	Sequence 31, App1
34	959.5	36.0	250 10	US-09-880-748-1952	Sequence 1952, Ap
35	959.5	36.0	250 15	US-10-293-418-1952	Sequence 1952, Ap
36	957.5	35.9	246 10	US-09-880-748-1268	Sequence 1268, Ap
37	957.5	35.9	246 15	US-10-293-418-1268	Sequence 1268, Ap
38	954	35.8	251 10	US-09-880-748-1238	Sequence 1238, Ap
39	954	35.8	251 15	US-10-293-418-1238	Sequence 1238, Ap
40	946	35.5	434 14	US-10-157-408-4	Sequence 4, App11
41	946	35.5	434 14	US-10-097-044A-4	Sequence 4, App11
42	945	35.4	448 16	US-10-769-247-4	Sequence 4, App11
43	945	35.4	448 14	US-10-024-329-32	Sequence 32, App1
44	940	35.2	370 9	US-09-759-841-6	Sequence 6, App11
45	940	35.2	433 17	US-10-872-198-129	Sequence 129, App

ALIGNMENTS

RESULT 1  
US-10-151-274-3  
; Sequence 3, Application US/10151274  
; Publication No. US20030064071A1  
; GENERAL INFORMATION:  
; APPLICANT: Littman, Dan R.  
; APPLICANT: Kwon, Douglas S.  
; APPLICANT: van Kooyk, Yvette  
; APPLICANT: Geiljenbeck, Theo  
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
; TITLE OF INVENTION: INVO  
; FILE REFERENCE: 1049-1-017  
; CURRENT APPLICATION NUMBER: US/10/151,274  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US/09/517,605  
; PRIOR FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; US-10-151-274-3  
Query Match 39.9%; Score 1064; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 7.4e-51;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVLQALALPAAATGGKVVYLGKGGDVEVLTCTASOKKSIOFHKNNOIK 60  
DB 1 MNRGVPFRHLLVLQALALPAAATGGKVVYLGKGGDVEVLTCTASOKKSIOFHKNNOIK 60  
QY 61 ILNQGSLFKGPKSLNDRADSRRSIMDQGNFPLIIKNLIKEDSDTYICEVEDOKEEVOL 120

```

Db      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208

```

```

RESULT 2
US-10-103-597A-39
; Sequence 39, Application US/10103597A
; Publication No. US20030096432A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; PRIOR FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-597A-39

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 3
US-10-188-444-39
; Sequence 39, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; PRIOR FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458

```

```

; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-188-444-39

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208

```

```

RESULT 4
US-10-207-655-170
; Sequence 170, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; PRIOR FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-170

Query Match      39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7.4e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Db      1  MNRGVPRRHLLVLQALPLPAATQGGKVVLGKKGDTVELTCTASQKKSIQFHWKNSNOIK 60
Qy      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Db      61  ILGNQGSFLLTKGSPSKLNDRAADSRSLMDQGNFPLIINKLKIESDITYICEVEDQKEEVOL 120
Qy      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Db      121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
Qy      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208
Db      181 TWTCTVLQONQKVEFKIDIVVLAFOKAS 208

RESULT 5
US-08-681-219-27
; Sequence 27, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; SIGNAL-TRANSDUCING PROTEINS AND THE GPCR

```

```

; TITLE OF INVENTION: (PDZ/DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Cooper & Dunham LLP
; STREET: 1185 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10036
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/681,219
; FILING DATE: 22-JUL-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: White, John P
; REGISTRATION NUMBER: 28,678
; REFERENCE/DOCKET NUMBER: 0575/48962/JPM/JKM
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 278-0400
; TELEFAX: (212) 391-0525
; INFORMATION FOR SEQ ID NO: 27:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 458 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-08-681-219-27

```

```

Query Match      39.8%; Score 1062; DB 8; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
DB 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 6
US-09-230-111C-25
; Sequence 25, Application US/09230111C
; Publication No. US20030203414A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; APPLICANT: Yanagisawa, Jun
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GLGF (PDZ/DHR)
; FILE REFERENCE: 48962-A-PCT-US
; CURRENT APPLICATION NUMBER: US/09/230,111C
; CURRENT FILING DATE: 1999-05-17
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 458
; TYPE: PRT

```

```

; ORGANISM: human
US-09-230-111C-25

```

```

Query Match      39.8%; Score 1062; DB 10; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
DB 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 7
US-10-092-138-25
; Sequence 25, Application US/10092138
; Publication No. US20030170723A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT
; CURRENT APPLICATION NUMBER: US/10/092,138
; CURRENT FILING DATE: 2002-09-06
; NUMBER OF SEQ ID NOS: 33
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 25
; LENGTH: 458
; TYPE: PRT
; ORGANISM: human
; US-10-092-138-25

```

```

Query Match      39.8%; Score 1062; DB 14; Length 458;
Best Local Similarity 99.5%; Pred. No. 9.5e-51;
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
DB 1 MNRGVPFRHLLVLTQALLPAATQGRKVVLGKGGDTVELTCTASQKKSIOFHKNSNQIK 60
QY 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
DB 61 ILNQGSPFLTKGSPSKINDRADSRSLMDQGNFPLIINKLIEBDDTYICEVEEQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
DB 121 LVFGLTANSDTHLLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180
QY 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208
DB 181 TWTCTVLQNGKVEFKIDIVLAFOKAS 208

```

```

RESULT 8
US-10-820-403-25
; Sequence 25, Application US/10820403
; Publication No. US20040229287A1
; GENERAL INFORMATION:
; APPLICANT: Sato, Taka-Aki
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION
; FILE REFERENCE: 65823/JPW/PT

```

CURRENT APPLICATION NUMBER: US/10/820,403  
CURRENT FILING DATE: 2004-04-08  
PRIOR APPLICATION NUMBER: US/10/092,138  
PRIOR FILING DATE: 2002-03-06  
NUMBER OF SEQ ID NOS: 34  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 25  
LENGTH: 458  
TYPE: PRT  
ORGANISM: human  
US-10-820-403-25

Query Match 39.8%; Score 1062; DB 16; Length 458;  
Best Local Similarity 99.5%; Pred. No. 9.5e-51;  
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAATGKRVVGGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
DB 1 MNRGVPFRHLVQLALPAATGKRVVGGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
QY 61 ILGNQGSFLTKGSKLNDRAISRSLWDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGSKLNDRAISRSLWDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVVLAFOKAS 208

## RESULT 9

US-09-939-537-29  
Sequence 29, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:

APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS

NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Clark & Elbling LLP  
STREET: 176 Federal Street  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbling, Karen L.  
REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>

INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 29:

US-09-939-537-29  
Query Match 39.7%; Score 1059; DB 10; Length 398;  
Best Local Similarity 99.5%; Pred. No. 1.2e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVQLALPAATGKRVVGGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
DB 1 MNRGVPFRHLVQLALPAATGKRVVGGKGDVVELTCTASQKKSIOFHWKNSNQIK 60  
QY 61 ILGNQGSFLTKGSKLNDRAISRSLWDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGSKLNDRAISRSLWDQGNFPLIINKLIKIDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLQGSQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVVLAFOKAS 208

## RESULT 10

US-10-157-408-1  
Sequence 1, Application US/10157408  
Publication No. US20030104535A1  
GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.  
Gregory, Timothy J.

TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:

ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco

STATE: California  
COUNTRY: USA  
ZIP: 94080

COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patIn (Genentech)

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/157,408  
FILING DATE: 28-May-2002  
CLASSIFICATION: 435

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329

FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-157-408-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1,2e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPRHLLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDEQKEEVOL 120  
DB 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDEQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208

RESULT 11  
US-10-097-044A-1  
Sequence 1, Application US/10097044A  
Publication No. US20030143220A1  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/097,044A  
FILING DATE: 28-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785

FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-097-044A-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1,2e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPRHLLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAATQGNKVVLGKKGDTVELTCTASQKKSIOFHWKNSNOIK 60  
QY 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDEQKEEVOL 120  
DB 61 ILNQGSLFKGSPKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDEQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDTHLLQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVQLNQKQKVEFKIDIVLAFOKAS 208

RESULT 12  
US-10-769-247-1  
Sequence 1, Application US/10769247  
Publication No. US20040197809A1  
GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/769,247  
FILING DATE: 30-Jan-2004  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777

FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-769-247-1

Query Match 39.7%; Score 1059; DB 16; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1,4e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLHLVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLHLVQLALLPAATGKNKVVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILNQGSLFKGSPSKNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
DB 61 ILNQGSLFKGSPSKNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180  
DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180  
QY 181 TWCTVVLQNKVKEFKIDIVLAFQKAS 208  
DB 181 TWCTVVLQNKVKEFKIDIVLAFQKAS 208  
RESULT 13  
US-09-891-119A-9  
Sequence 9, Application US/09891119A  
Publication No. US20040013683A1  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
FILE REFERENCE: 24577-CY-B  
CURRENT APPLICATION NUMBER: US/09/891,119A  
CURRENT FILING DATE: 2001-06-25  
NUMBER OF SEQ ID NOS: 22  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 9  
LENGTH: 457  
TYPE: PRT  
ORGANISM: human  
US-09-891-119A-9

Query Match 39.7%; Score 1059; DB 11; Length 457;  
Best Local Similarity 99.5%; Pred. No. 1,4e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLHLVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLHLVQLALLPAATGKNKVVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILNQGSLFKGSPSKNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
DB 61 ILNQGSLFKGSPSKNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180

DB 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLISVQLELQDSG 180  
QY 181 TWCTVVLQNKVKEFKIDIVLAFQKAS 208  
DB 181 TWCTVVLQNKVKEFKIDIVLAFQKAS 208

RESULT 14  
US-09-939-537-5  
Sequence 5, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanue, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-939-537-5

Query Match 39.7%; Score 1059; DB 10; Length 462;  
Best Local Similarity 99.5%; Pred. No. 1,4e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLHLVQLALLPAATGKKVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLHLVQLALLPAATGKNKVVVLGKKGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILNQGSLFKGSPSKNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEYOL 120

Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPIIRKLIKIEDSDTYICEVEDQKEVQL 120  
Qy 121 LVFGLTANSPDTHLLOQGSLLTLTLESPPGSSPSVQCRRSPRGKNIQGGKTLTSSVQLLEQDSG 180  
Db 121 LVFGLTANSPDTHLLOQGSLLTLTLESPPGSSPSVQCRRSPRGKNIQGGKTLTSSVQLLEQDSG 180  
Qy 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208

## RESULT 15

US-09-243-008-5  
; Sequence 5, Application us/09243008  
; Publication No. US20040005334A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Redirection of Cellular Immunity by  
; Receptor Chimeras  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,008  
; FILING DATE: 02-Feb-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,176  
; FILING DATE: SEPTEMBER 11, 1995  
; APPLICATION NUMBER: 08/203,866  
; FILING DATE: February 28, 1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: March 6, 1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karen F. Iech, Ph.D.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/270001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 462 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
US-09-243-008-5

Query Match 39.7%; Score 1059; DB 11; Length 462;  
Best Local Similarity 99.5%; Pred. No. 1.4e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 MNRGVPFRHLVQLALPAAATQKGVLAGKGDIVELTCTASQKKSIOFHWKNSNOIK 60  
Db 1 MNRGVPFRHLVQLALPAAATQKGVLAGKGDIVELTCTASQKKSIOFHWKNSNOIK 60  
Qy 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPIIRKLIKIEDSDTYICEVEDQKEVQL 120  
Db 61 ILGNQGSFLTKGSPSKLNDRADSRSLMDQGNFPIIRKLIKIEDSDTYICEVEDQKEVQL 120

Qy 121 LVFGLTANSPDTHLLOQGSLLTLTLESPPGSSPSVQCRRSPRGKNIQGGKTLTSSVQLLEQDSG 180  
Db 121 LVFGLTANSPDTHLLOQGSLLTLTLESPPGSSPSVQCRRSPRGKNIQGGKTLTSSVQLLEQDSG 180  
Qy 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLQONOKKVEFKIDIVLAFQKAS 208

Search completed: September 20, 2005, 16:27:40  
Job time : 487 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 20, 2005, 13:43:29 ; Search time 34 Seconds

(without alignments)  
1437.592 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRRLILVLQALLP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

283416

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	1 RWHUT4	T-cell surface gly
2	916	34.3	432	1 RWCZT4	T-cell surface gly
3	831	31.1	432	1 RWMQT4	T-cell surface gly
4	750	28.1	268	2 A56446	Ig heavy chain V r
5	668	25.0	249	2 S41374	single chain Fv an
6	618.5	23.2	459	2 A46254	CD4 precursor - ra
7	616	23.1	233	2 JC5322	p3 specific singl
8	596	22.3	432	2 S30193	T-cell surface gly
9	508.5	19.1	457	2 A27449	T-cell surface gly
10	507	19.0	128	2 S40343	Ig kappa chain V-J
11	498	18.7	107	2 S34005	Ig kappa chain V r
12	498	18.7	117	2 S40362	Ig kappa chain - h
13	498	18.7	457	1 RWMST4	T-cell surface gly
14	485.5	18.2	116	2 B26555	Ig kappa chain V-I
15	485	18.2	123	2 S40378	Ig kappa chain - h
16	481.5	18.0	131	2 S40358	Ig kappa chain - h
17	480.5	18.0	109	1 K3HUPM	Ig kappa chain V-I
18	475.5	17.8	215	2 JE0244	Ig kappa chain NIG
19	466	17.5	111	2 S23628	Ig kappa chain V r
20	460	17.2	108	2 G44151	Ig kappa chain V r
21	458	17.2	128	2 A56701	Ig kappa chain V r
22	456.5	17.1	114	2 S54905	Ig kappa chain V r
23	456	17.1	128	2 S40345	Ig kappa chain V-J
24	455	17.1	128	2 S40375	Ig kappa chain V-J
25	455	17.1	129	2 S29627	Ig kappa chain V r
26	454	17.0	115	1 K3HUCT	Ig kappa chain pre
27	453	17.0	98	2 I30608	Ig kappa chain V-I
28	453	17.0	115	2 A30553	Ig kappa chain pre
29	453	17.0	144	2 PL0106	Ig kappa chain pre

#### ALIGNMENTS

##### RESULT 1

RWHUT4 T-cell surface glycoprotein CD4 precursor [validated] - human

N:Alternate names: T-cell surface antigen T4/Len 3

C:Species: Homo sapiens (man)

C>Date: 28-May-1986 #sequence revision 31-Dec-1988 #text change 09-Jul-2004

C:Accession: A90872; A32722; A34194; A53287; I54176; I54297; A02109; A30039

R:Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Cheese, L.; Axel, R.

Cell 42, 93-104, 1985

A:Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface prot

A:Reference number: A90872; WUID:85254948; PMID:2990730

A:Accession: A90872

A:Molecule type: mRNA

A:Residues: 1-25, 'N', 27-458 <MAD>

A:Cross-references: UNIPROT:P01730

A:Experimental source: clone p74B

R:Littman, D.R.; Maddon, P.J.; Axel, R.

Cell 55, 541, 1988

A:Title: Corrected CD4 sequence.

A:Reference number: A90807; WUID:89028665; PMID:3263213

A:Contents: annotation; revision to residue 26

R:Cammerlin, D.; Seed, B.

Cell 60, 747-754, 1990

A:Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi

A:Reference number: A32722; WUID:90182664; PMID:2107024

A:Accession: A32722

A:Molecule type: mRNA

A:Residues: 26-426, 428-458 <CAM>

A:Status: nucleic acid sequence not shown; not compared with conceptual translation

A:Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recept

A:Reference number: A34194; WUID:90078232; PMID:2592374

A:Contents: disulfide bonds; carbohydrate-binding sites

A:Accession: A34194

A:Molecule type: protein

A:Residues: 26-354 <CAR>

R:Lederman, S.; Demartino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A.

Mol. Immunol. 28, 1171-1181, 1991

A:Title: A single amino acid substitution in a common African allele of the CD4 molecule

A:Reference number: A53287; WUID:92072595; PMID:1961196

A:Accession: A53287

A:Status: not compared with conceptual translation

A:Molecule type: mRNA

A:Residues: 250-264, 'W', 266-280 <LED>

A>Note: sequence extracted from NCBI backbone (NCBIP.68249)

R:Edwards, M.C.; Gibbs, R.A.

Genomics 14, 590-597, 1992

A:Title: A human dimorphism resulting from loss of an Alu.

A:Reference number: I54176; WUID:93052367; PMID:1330888

A:Accession: I54176

A:Status: translated from GB/EMBL/DBJ

A:Molecule type: DNA  
 A:Residues: 1-72 <RES>  
 A:Cross-references: GB:U47924; GB:M86525; GB:U72506; NID:G1633547; PIDN:AAB51309.1; PID:  
 R:Hodge, T.W.; Sasso, D.R.; McConugal, J.S.  
 Hum. Immunol. 30, 99-104, 1991  
 A>Title: Humans with OK4-epitope deficiency have a single nucleotide base change in the  
 A:Reference number: I54297; MUID:91216786; PMID:11708753  
 A:Accession: I54297  
 A:Status: translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1264, 'W', 266-458 <RES>  
 A:Cross-references: GB:M35160; NID:G179143; PIDN:AAA16069.1; PID:G179144  
 C:Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) recep  
 C:Genetics:  
 A:Gene: GDB:CD4  
 A:Cross-references: GDB:119767; OMIM:186940  
 A:Map position: 12pter-12p12  
 A:Features: 16/3  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:26-48/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
 F:34-111/Domain: immunoglobulin homology <IM1>  
 F:136-186/Domain: immunoglobulin homology #status atypical <IM2>  
 F:216-299/Domain: immunoglobulin homology <IM3>  
 F:321-372/Domain: immunoglobulin homology <IM4>  
 F:397-420/Domain: transmembrane #status predicted <TM>  
 F:421-458/Domain: intracellular #status predicted <INT>  
 F:41-109,155-184,328-370/Disulfide bonds: #status experimental  
 F:266,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1e-48;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVQLALPAATGGKRVYLGKSDTYELCTASQKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLIVQLALPAATGGKRVYLGKSDTYELCTASQKKSIOFHMKNNOIK 60  
 QY 61 ILNQGSEFLTKGSPKLNDRADSRSLMDQGNPFLIKLIEBDDTYICEVEDQKEVQL 120  
 DB 61 ILNQGSEFLTKGSPKLNDRADSRSLMDQGNPFLIKLIEBDDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSLTLTLESPGSSPVOCRSRPNKIQGKTLVSQLELDGSG 180  
 DB 121 LVFGLTANSDTHLLQGSLTLTLESPGSSPVOCRSRPNKIQGKTLVSQLELDGSG 180  
 QY 181 TWTCITVLQNKVYEFKIDIVVLAFOKAS 208  
 DB 181 TWTCITVLQNKVYEFKIDIVVLAFOKAS 208

RESULT 2  
 RMC274  
 T-cell surface glycoprotein CD4 - chimpanzee  
 N:Alternate names: T-cell surface antigen T4/Leu 3  
 C:Species: Pan troglodytes (chimpanzee)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C:Accession: B32722; A46534  
 R:Cametini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A>Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A:Reference number: A32722; MUID:90182664; PMID:2107024  
 A:Accession: B32722  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <CAM>  
 A:Cross-references: UNIPROT:P16004; GB:M31135  
 R:Fomsgaard, A.; Hirsch, V.M.; Johnson, P.R.  
 Eur. J. Immunol. 22, 2973-2981, 1992  
 A>Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular recep  
 A:Reference number: A46534; MUID:93049640; PMID:11455921  
 A:Accession: A46534  
 A:Status: not compared with conceptual translation

A:Molecule type: mRNA  
 A:Residues: 3-399 <FOM>  
 A>Note: sequence extracted from NCBI backbone (NCBI:P118332)  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F:1-371/Domain: extracellular #status predicted <EXT>  
 F:9-86/Domain: immunoglobulin homology <IM1>  
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F:191-274/Domain: immunoglobulin homology <IM3>  
 F:286-347/Domain: immunoglobulin homology <IM4>  
 F:312-395/Domain: transmembrane #status predicted <TM>  
 F:396-432/Domain: intracellular #status predicted <INT>  
 F:16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.3%; Score 916; DB 1; Length 432;  
 Best Local Similarity 97.8%; Pred. No. 4.9e-41;  
 Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 26 KKVVLRKKGDTVELCTASQKKSIOFHMKNNOIKILNQGSEFLTKGSPKLNDRADSR 85  
 DB 1 KKVVLRKKGDTVELCTASQKKSIOFHMKNNOIKILNQGSEFLTKGSPKLNDRADSR 85  
 QY 86 LMDQGNPFLIKLIEBDDTYICEVEDQKEVQLVFGLTANSDTHLLQGSLTLTLES 145  
 DB 61 LMDQGNPFLIKLIEBDDTYICEVEDQKEVQLVFGLTANSDTHLLQGSLTLTLES 120  
 QY 146 PPGSSSPVOCRSRPNKIQGKTLVSQLELDGSGTWTCITVLQNKVYEFKIDIVVLAQ 205  
 DB 121 PPGSSSPVOCRSRPNKIQGKTLVSQLELDGSGTWTCITVLQNKVYEFKIDIVVLAQ 180  
 QY 206 KAS 208  
 DB 181 KAS 183

RESULT 3  
 RMC274  
 T-cell surface glycoprotein CD4 - rhesus macaque  
 N:Alternate names: T-cell surface antigen T4/Leu 3  
 C:Species: Macaca mulatta (rhesus macaque)  
 C>Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C:Accession: C32722  
 R:Cametini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A>Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A:Reference number: A32722; MUID:90182664; PMID:2107024  
 A:Accession: C32722  
 A:Molecule type: mRNA  
 A:Residues: 1-432 <CAM>  
 A:Cross-references: GB:M31134  
 C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C:Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F:1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F:1-371/Domain: extracellular #status predicted <EXT>  
 F:9-86/Domain: immunoglobulin homology <IM1>  
 F:111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F:180-293/Domain: immunoglobulin homology <IM3>  
 F:286-347/Domain: immunoglobulin homology <IM4>  
 F:312-395/Domain: transmembrane #status predicted <TM>  
 F:396-432/Domain: intracellular #status predicted <INT>  
 F:16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F:271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.1%; Score 831; DB 1; Length 432;  
 Best Local Similarity 88.0%; Pred. No. 1.3e-36;  
 Matches 161; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 26 KKVVLRKKGDTVELCTASQKKSIOFHMKNNOIKILNQGSEFLTKGSPKLNDRADSR 85  
 DB 1 KKVVLRKKGDTVELCTASQKKSIOFHMKNNOIKILNQGSEFLTKGSPKLNDRADSR 85

Db 1 KKVLLGKKGDVTELTCTASOKKNTQFHWKNSNOIKIIGIQLFLYTKGPSKLSDRADSRKS 60  
 QY 86 LMDGNFPLIITKNIKIEDSDTYICEVEDQKEEVOLLVGLFANSDTHLLOQSLLTLTES 145  
 Db 61 LMDGSCSMIITKNIKIEDSDTYICEVENKKEVELVFGFLFANSDTHLBSQSLLTLTES 120  
 QY 146 PGGSSPVOCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNKVFEKIDIVLAFQ 205  
 Db 121 PGGSSPVCKCRSPGKNIQGGRTISVQLERQDSGTWCTVLSQDKVFEKIDIVLAFQ 180  
 QY 206 KAS 208  
 Db 181 KAS 183

## RESULT 4

A56446  
 Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
 C/Species: Mus musculus (house mouse)  
 C/Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
 C/Accession: A56446  
 R/Teng, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
 J. Biol. Chem. 270, 7829-7835, 1995  
 A/Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical  
 A/Reference number: A56446; PMID:95229583; PMID:7713873  
 A/Accession: A56446  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-268 <TAN>  
 A/Cross-references: GB:U20617  
 C/Keywords: heterotrimer; immunoglobulin

Query Match 28.1%; Score 750; DB 2; Length 268;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-32;  
 Matches 150; Conservative 39; Mismatches 57; Indels 24; Gaps 7;  
 QY 243 SOVQLLESGAEVKKRGSSVVKSCAKSG---DTEIRYSFTVWROAPQGLEMMGRITITL 298  
 Db 2 AQVVLQSSGAEIVKRGASVYKLSCTTSGFNIMDTYMH---WVKRPQGLEWIGRIAPAN 57  
 QY 299 DVAAHAPHLQGRVITTDKSTSTVYLELRNRSDDTAAYFCAGYEGEADGEYDNNQFL 358  
 Db 58 GITKCDKPFQKATITADTSNTAYVLQSLTSSEDTAVYCASYY---LTRYEN--- 108  
 QY 359 KHWQGTLYVTSSGGSGSGSGGSELELTQSPATLSVPERATLSRAESVSSD 418  
 Db 109 -YMQQGTIVYVSSGGSGSGGSDIELTQSPAIMSASLDEKVTMSCRASSVNF- 166  
 QY 419 LAWYQKPGQAPRLIYGASTRATGVPARFSGSGSGAEFTLTISLQSEDPAYVYCOQYN 478  
 Db 167 IYMQQSDASPKLWVYTHSLPRGVPARFSGSGSGNSYSLTISMEGEDATYVYCOQFT 226  
 QY 479 NMPRYTFQGTTRLEIKLVPRSGHHHHH 508  
 Db 227 SSP--FTFGSGTKLEIK---RSAHHHHH 251

## RESULT 5

S41374  
 single chain Fv antibody - mouse  
 C/Species: Mus musculus (house mouse)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
 C/Accession: S41374  
 R/Altshenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.  
 submitted to the EMBL Data Library, January 1994  
 A/Description: Construction and functional characterization of a single chain Fv antibody  
 A/Reference number: S41374  
 A/Accession: S41374  
 A/Status: preliminary  
 A/Molecule type: DNA  
 A/Residues: 1-249 <ART>  
 A/Cross-references: EMBL:Z29480

Query Match 25.0%; Score 668; DB 2; Length 249;  
 Best Local Similarity 52.9%; Pred. No. 2.5e-28;  
 Matches 136; Conservative 40; Mismatches 65; Indels 16; Gaps 4;  
 QY 244 QVQLLESGAEVKKRGSSVVKSCAKSGDTFRISFTWYRQAPGGGLEWGGIITILDVAHY 303  
 Db 1 QVQLQSGAEIVKRGASVYKLSCTASGFNFQDVIHWKQRPKGLWIAIAPASGVNKY 60  
 QY 304 APHLQGRVITTDKSTSTVYLELRNRSDDTAAYFCAGYEGEADGEYDNNQFLKHWGQ 363  
 Db 61 VPRFQDATTATTSNTAYVLQSLTSSEDTAVYCA-----RDTLYTSLGK---WQ 111  
 QY 364 GTLVVTSGGSGSGSGSGGSELELTQSPATLSVPERATLSGRASESV-----SSD 418  
 Db 112 GSTVYVSSRGSGSGSGGSDIELTQSPVAVVPGSSVSISSRSSSLYSQSDSY 171  
 QY 419 LAWYQKPGQAPRLIYGASTRATGVPARFSGSGSGAEFTLTISLQSEDPAYVYCOQYN 478  
 Db 172 LFWFLQRPQSPQLLIYRMSNLASGVDPDRFSGSGSGFTLRISRYEADWGYVYCWQHR 231  
 QY 479 NMPRYTFQGTTRLEIK 495  
 Db 232 EYP--LTFGAGTYLEIK 246

## RESULT 6

A46254  
 CD4 precursor - rabbit  
 C/Species: Oryctolagus cuniculus (domestic rabbit)  
 C/Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C/Accession: A46254  
 R/Hague, B.F.; Sawaadikosol, S.; Brown, T.J.; Lee, K.; Recker, D.P.; Kindt, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
 A/Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency virus  
 A/Reference number: A46254; PMID:92390370; PMID:1518821  
 A/Accession: A46254  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-459 <RAG>  
 A/Cross-references: UNIPROT:P46630; GB:M92840; NID:g164871; PIDN:AAA31198.1; PID:g164872  
 A/Note: sequence extracted from NCBI backbone (NCBI:112732, NCBI:P:112733)  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 23.2%; Score 618.5; DB 2; Length 459;  
 Best Local Similarity 59.6%; Pred. No. 1.7e-25;  
 Matches 127; Conservative 34; Mismatches 47; Indels 5; Gaps 2;  
 QY 1 MNRGVPRHLLVYLALPAATQGGKVVYGGKDTVELTCTASOKKSIQFHKNSNOIK 60  
 Db 1 MNRRIYFQCLLVLPALPAATWGTIVYRGKAGALVELPCOSSQKNSVFMKHAQVK 60  
 QY 61 ILNQG-----SPLTKGPSKLNDRADSRSLMDGNFPLIITKNIKIEDSDTYICEVEDQKE 116  
 Db 61 ILNQGSSSSSPWLKNGSPLSNVRVESKKNWMDGSPFLVYKIDLRMDSGTYICEVEDQKM 120  
 QY 117 EVQLVFPGLTANSDTHLLOQSITLTLESPPGSSPSVQCRSPGKNIQGGKTLVSQLEL 176  
 Db 121 EVELVFRLLTANSDTHLHQSLTLTLEGSSVSPVQWKSPPNKIIETGPTCSMPKRL 180  
 QY 177 QDSGTWCTV-LQNKVFEKIDIVLAFQKAS 208  
 Db 181 QDSGTWSCHSPQDKLELDIKIIVLGFPKAS 213

## RESULT 7

JC5322  
 p53 specific single-chain antibody Pab421 - human  
 C/Species: Homo sapiens (man)  
 C/Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
 C/Accession: JC5322  
 R/Jannot, C.B.; Hynes, N.E.  
 Biochem. Biophys. Res. Commun. 230, 242-246, 1997

A>Title: Characterization of acFv-421, a single-chain antibody targeted to p53.  
 A/Reference number: JCS322; MUID:97168950; PMID:9016757  
 A/Accession: JCS322  
 A/Molecule type: mRNA  
 A/Residues: 1-233 <JAN>  
 A:Experimental source: hybridoma cell  
 C/Comment: This protein specifically binds the tumor suppressor protein p53. It restores

Query Match 23.1%; Score 616; DB 2; Length 233;  
 Best Local Similarity 52.6%; Pred. No. 1.2e-25;  
 Matches 132; Conservative 32; Mismatches 63; Indels 24; Gaps 4;

QY 249 ESGAEVKKPPGSSVSVSCASDPTIRSFYVWRAPQSGLEFMKRITITLDVAHYAHLQ 308  
 DB 2 ESGAEVRSQASVYLSCTTSQFNINDYMHVKKRPPQSGLEWGRIPENDGDMTRSSG 61  
 QY 309 GRVITTDKSTSYLYELRNRSDDTVYFCAGYEEBAGEADGVDNNGLKMGQGTLYT 368  
 DB 62 VKATMTADTSNTAYLQLSSLTSEDVAYYC-----NAGMDYMGQGTITVT 106  
 QY 369 VTSGGSGSGGSGGSGGSELELTQSPATLSVSPERATLSCRASESYS---SDLAWYQQ 424  
 DB 107 VSSGSGSGGSGRASGSGSDIELTQSPASLAVSLGQRAITISCRASKSVSTSGSYMHNNQ 166  
 QY 425 KPGQAPRLIYGASTRATGVPARFSGSGSGAEFTLTSSLOSEDFANYCCQYNNMPRY 484  
 DB 167 KPGQAPRLIYLVNLSGVPARFSGSGSGDTFTLNIHPVEEDATYCOHTR---EL 222  
 QY 485 TFGQ-GTRLEI 494  
 DB 223 TRSEGTRLEI 233

## RESULT 8

S30193  
 T-cell surface glycoprotein CD4 - dog  
 C/Species: Canis lupus familiaris (dog)  
 C/Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 21-Jul-2000  
 C/Accession: S30193  
 R/Mide, K.F.; Conner, G.E.; Mintz, D.H.; Alejandro, R.  
 Biochim. Biophys. Acta 1172, 315-318, 1993  
 A>Title: Primary structure of the canine CD4 antigen.  
 A/Reference number: S30193; MUID:93192324; PMID:7916632  
 A/Accession: S30193  
 A/Status: preliminary  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <MIL>  
 A:Cross-references: EMBL:X68565; NID:9288652; PIDD:CA837664.1; PIDD:94467377  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: glycoprotein  
 F/202-311/Domain: immunoglobulin homology <IMM>

Query Match 22.3%; Score 596; DB 2; Length 432;  
 Best Local Similarity 56.3%; Pred. No. 2.4e-24;  
 Matches 116; Conservative 38; Mismatches 42; Indels 10; Gaps 2;

QY 12 LVLTQALLPAPATQKRVVLRGKGDVTELTCTASQKSIQPHMKNSNOIKILGNQSFYTK 71  
 DB 1 LMTQVLWLPATVPREVVLGKADAVELPCQTSQKNIHFMWRSSWVQILGNQSFYTV 60  
 QY 72 GPEGLNDRADSRRLMOQGNFPLIILKULKIEDSTTYICEVEDQKEEVQLVFGILTA--- 127  
 DB 61 GSSRLKRRVSKRLMVGSPFLVYIKDLVADSGIYCDT-DKQOEVELLVFNLTAKWDS 119  
 QY 128 -----NSDTLHLOGSLTLTLESPGSSPSVQCRSPGKNIQSGKTLVSQLELDQSGTW 182  
 DB 120 GSSSGSSNIRLQOQQLTLTENPSSGSSPSVQMGKPNKSGHGONLSLWPELDGQIW 179  
 QY 183 TCTVLQNKVVEFKIDIVVLAFOKAS 208  
 DB 180 TCTISQSKTVEFNIVVLAFOKAS 205

## RESULT 9

A27449  
 T-cell surface glycoprotein CD4 precursor - rat  
 N/Alternate names: W3/25 antigen  
 C/Species: Rattus norvegicus (Norway rat)  
 C/Date: 21-May-1988 #sequence\_revision 21-May-1988 #text\_change 09-Jul-2004  
 C/Accession: A27449; A35433  
 R/Clark, S.J.; Jelfries, W.A.; Barclay, A.N.; Gagnon, J.; Williams, A.F.  
 Proc. Natl. Acad. Sci. U.S.A. 84, 1649-1653, 1987  
 A>Title: Peptide and nucleotide sequences of rat CD4 (W3/25) antigen: evidence for derive  
 A/Reference number: A27449; MUID:87155535; PMID:3104900  
 A/Accession: A27449  
 A/Molecule type: mRNA  
 A/Residues: 1-457 <CLA>  
 A:Cross-references: UNIPROT:P05540; GB:M15768; NID:9203387; PIDD:AAA40901.1; PIDD:9203388  
 R/Davis, S.J.; Ward, H.A.; Puklavec, M.J.; Willis, A.C.; Williams, A.F.; Barclay, A.N.  
 J. Biol. Chem. 265, 10410-10418, 1990  
 A>Title: High level expression in Chinese hamster ovary cells of soluble forms of CD4 T  
 A/Reference number: A35433; MUID:90285164; PMID:2113054  
 A/Contents: annotation  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: glycoprotein; membrane protein; surface antigen  
 F/219-300/Domain: immunoglobulin homology <IMM>

Query Match 19.1%; Score 508.5; DB 2; Length 457;  
 Best Local Similarity 49.3%; Pred. No. 9.2e-20;  
 Matches 104; Conservative 34; Mismatches 70; Indels 3; Gaps 2;

QY 1 MNRGVPFRHL--LLVQLALLPAPATQKRVVLRGKGDVTELTCTASQKSIQPHMKNSQ 58  
 DB 1 MCRGSPFRHLPLLLQLSLVLTQKTVVLRGKSGSELPCESTRSRASAFAMKSDQ 60  
 QY 59 IKILGNQSFYTKGPKLNDADSRRLWDQGNFPLIILKULKIEDSTTYICEVEDQKEEV 118  
 DB 61 KTLIGYKNKLLIGSLVSRFDSRKNAWERSGFPILIKLRNEDSQTYVCELENKEEV 120  
 QY 119 QLVFGLTANSPTHLQSGSLTTLTLES-PPGSSPSVQCRSPGKNIQSGKTLVSQLEIQ 177  
 DB 121 ELWVFVTFNPGFRLLQGGSLTLTLDSPKVPSPPIECRKSNIWDSKAFSTHSLRIQ 180  
 QY 178 DSGTWTCYVLQNKVVEFKIDIVVLAFOKAS 208  
 DB 181 DSGTWTCYVLQNKVVEFKIDIVVLAFOKAS 211

## RESULT 10

S40343  
 Ig kappa chain V-J region - human  
 C/Species: Homo sapiens (man)  
 C/Date: 19-May-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
 C/Accession: S40343  
 R/Klein, R.; Jaenichen, R.; Zachau, H.G.  
 Eur. J. Immunol. 23, 3248-3271, 1993  
 A>Title: Expressed human immunoglobulin chi genes and their hypermutation.  
 A/Reference number: S40343; MUID:94080891; PMID:8258341  
 A/Accession: S40343  
 A/Status: preliminary; translation not shown  
 A/Molecule type: mRNA  
 A/Residues: 1-128 <KLE>  
 A:Cross-references: EMBL:X72453; NID:9441374; PIDD:CAA51121.1; PIDD:9441375  
 C/Superfamily: immunoglobulin V region; immunoglobulin homology  
 C/Keywords: heterotrimer; immunoglobulin  
 F/36-110/Domain: immunoglobulin homology <IMM>

Query Match 19.0%; Score 507; DB 2; Length 128;  
 Best Local Similarity 89.0%; Pred. No. 3.3e-20;  
 Matches 97; Conservative 7; Mismatches 3; Indels 2; Gaps 1;

QY 387 ELELTQSPATLSVSPERATLSCRASESYSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 21 EIVMTQSPATLSVSPERATLSCRAQSQSVSNLAWYQKPGQAPRLIYGASTRATGVA 80  
 QY 447 RFGSGSGAEFTLTSSLOSEDFANYCCQYNNMPRYTFGQTRLEIK 495



	Query Match	18.7%	Score 498;	DB 1;	Length 457;
	Best Local Similarity	51.6%	Pred. No. 3.3e-19;		
	Matches	110;	Conservative	35;	Mismatches 62; Indels 6; Gaps 5
OY	1	MNRGVPRRH-LLVLTQALLPAATOGKVVLAGKKDVTLETCVTSOQKSIOFHMKNNSNOI	59		
Db	1	MCRAISLRLLILLLOLSQLAAVTOGKVLIVGEKESEBELPCSSOKKITFTWTFYSOR	60		
OY	60	KILINOG-SPLTKG-PSKLNDRADRSRLMDQGFPILINKLIKEDSDTYICEVDQKE	116		
Db	61	KILOHGKGAVIRGGSPSQ-DREPSKGNHEKSFPLLINKLMEDSGTYICELENRKE	119		
OY	117	EVOALLVEGLIANSDTHLIQOSGLTITLES-PPGSSPVQCSPPRGKNTIQGKTTLSVQLE	175		
Db	120	EVELMVFKEVTPSPGTSLMQOSGLTITLTDLSNKSANPLETECKHKKGKVSVGSKVLMSMLR	179		
OY	176	LDSGDTWTCTVLONOKKVEFKIDIVVLAFOKAS	208		
Db	180	VQSDSDFNNCTVTLDOCKGMFGMTLTSVLGFQSTIA	212		

```

RESULT 14
B26555
IG kappa chain V-III region (ger) - human
C:Species: Homo sapiens (man)
C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 21-Jan-2000
C:Accession: B26555
R:Midgaugh, C.R.; Litman, G.W.
J. Biol. Chem. 262, 3671-3673, 1987
A:Title: Atypical glycosylation of an IGG monoclonal cryoimmunoglobulin.
A:Reference number: A92630, MUID:87137666, PMID:3102493
A:Accession: B26555
A:Molecule type: protein
A:Residues: 1-116 <MID>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotetramer; immunoglobulin
E:16-90/Domain: immunoglobulin homology <IWM>

```

Query Match	18.2%	Score 485.5;	DB 2;	Length 116;
Best Local Similarity	85.3%;	Pred. No. 4e-19;		
Matches	93;	Conservative	9;	Mismatches 6; Indels 1; Gaps 1

  

QY	387	ELELTGSPATISVSGERATITSCRSRSESVSIIAWQKPGQAPRLIIYGASTRATGPA	446
Db	1	DIWVTQVPATISVSPGEGATITSCRSRSGSISSIIANLYQKPGQAPRLIIYMASTRATGIPA	60
QY	447	RFSSGSGCAEFTLTIIISLQSEDFAVYYCOQINNNPRTYFGQGRLEIK	495
Db	61	RFSSGSGGTEFTLTIIISLQSEDFAVYYCOQYDDWDP-ITFGQGRLEIK	108

RESULT 15  
S40378  
IG kappa chain - human  
C:Species: Homo sapiens (man)  
C:Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C:Accession: S40378  
R:Klein, R.; Jaenichen, R.; Zachau, H.G.  
Eut. J. Immunol. 23, 348-3271, 1993  
A:Title: Expressed human immunoglobulin chi genes and their hypermutation.  
A:Reference number: S40312; PMID:9408091; PMID:9285341  
A:Accession: S40378  
A:Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-123 <RL>  
A:Cross-references: EMBL:Y72488; NID:g441444; PDB:CAA51156.1; PDB:g441445  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
;25-103/Domain: immunoglobulin homology <IM>

Query Match	18.2%;	Score 485;	DB 2;	Length 123;
Best Local Similarity	80.4%;	Pred. No. 4.4e-19;		
Matches 90; Conservative	10;	Mismatches 12;	Indels 0;	Gaps 0;

```

QY      384  GGGELELLTOSPATLIVSVSPERATLISCRASIESVSSPLMAYCOQKPCQARLLLYVGASTATG  444
      11  GICEIWMTOSPATLIVSPEDTATLISCRASOGSSNLAMVQHRPCQAPRLLLYVGASTRAAG  70
Db
QY      444  VPARFSGSSGABFTLITSLILOSEDFAYAYCOQVNNMPRTYFQCGRLREIX  495
      71  IPRRFSGSSGTEFTLITSLIQSNAFAYAYCOQYIDMPEWMTYFQCGRLDIX  122
Db

Search completed: September 20, 2005, 16:10:09
Job time : 36 secs

```

Search completed: September 20, 2005, 16:10:05  
Job time : 36 secs

Job time : 36 secs

Search completed: September 20, 2005, 16:10:03  
Job time: 0.36 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 20, 2005, 12:03:07 ; Search time 137 Seconds  
(without alignments)  
1898.805 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRRLVLTQLALP.....GTRLEIKIVPRSGHHHHH 508

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	458	1	CD4_HUMAN
2	1040	39.0	458	1	CD4_PANTR
3	953	35.7	458	1	CD4_MACRA
4	951	35.6	458	1	CD4_MACRU
5	945	35.4	458	1	CD4_MACMU
6	944	35.4	458	1	CD4_MACNE
7	926	34.7	458	1	CD4_CERAB
8	824.5	30.9	244	2	065ZC8
9	820	30.7	397	1	CD4_CERPO
10	819	30.7	397	1	CD4_ERIPA
11	803	30.1	397	2	009261
12	803	30.1	397	2	009262
13	803	30.1	397	2	009263
14	803	30.1	397	2	009264
15	799	29.9	397	2	009265
16	799	29.9	397	2	009266
17	778.5	29.2	457	2	08H2H8
18	777.5	29.1	457	2	08H2H7
19	774.5	29.0	457	1	CD4_SAISC
20	766	28.7	255	2	06KB05
21	760.5	28.5	240	2	065ZC9
22	751.5	28.2	487	2	065ZL2
23	750	28.1	487	2	065ZL1
24	734	27.5	298	2	09QYR0
25	685.5	25.7	241	2	09QYR0
26	651	24.4	248	2	065Z07
27	632	23.7	463	1	CD4_CANFA
28	630	23.6	455	2	0710E2
29	625	23.4	455	2	09X578
30	618.5	23.2	459	1	CD4_RABIT
31	612	22.9	218	2	0925S1

32	596	22.3	432	2	06LBU1
33	590.5	22.1	474	2	P79355
34	585	21.9	457	2	06GYR3
35	563	21.1	406	2	06R3N3
36	563	21.1	447	2	06R3N4
37	508.5	19.1	457	1	CD4_RAT
38	505.5	18.9	129	1	KV3H_HUMAN
39	498	18.7	457	1	CD4_MOUSE
40	490	18.4	457	2	061396
41	485.5	18.2	109	2	09UL85
42	480.5	18.0	109	1	KV3P_HUMAN
43	480	18.0	108	2	09UL83
44	472.5	17.7	205	2	06GMW0
45	459.5	17.2	433	2	055054

## ALIGNMENTS

RESULT 1	CD4_HUMAN	STANDARD;	PRT;	458 AA.
AC	P01730;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	25-JAN-2005 (Rel. 46, Last annotation update)			
DE	T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).			
GN	Name=CD4;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=85254948; PubMed=2990730;			
RA	Maddon P.J., Littman D.R., Godfrey M., Maddon D.E., Chess L., Axel R.;			
RT	"The isolation and nucleotide sequence of a cDNA encoding the T cell surface protein T4: a new member of the immunoglobulin gene family.";			
RL	Cell 42:93-104(1985).			
RN	[2]			
RP	REVISION TO 26.			
RX	MEDLINE=89028665; PubMed=3263213; DOI=10.1016/0092-8674(88)90211-5;			
RA	Littman D.R., Maddon P.J., Axel R.;			
RT	"Corrected CD4 sequence.";			
RL	Cell 55:541-541(1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=96303695; PubMed=8723724;			
RA	Alasari-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,			
RT	"A gene-rich cluster between the CD4 and triosephosphate isomerase genes at human chromosome 12p13.";			
RL	Genome Res. 6:314-326(1996).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT TRP-265.			
RX	MEDLINE=91216786; PubMed=1708753; DOI=10.1016/0198-8859(91)90077-M;			
RA	Hodge T.W., Saaso D.R., McDougal J.S.;			
RT	"Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240.";			
RL	Hum. Immunol. 30:99-104(1991).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Pancreeas;			
RX	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RA	Krausner R.D., Collins F.S., Wagner L., Shennan C.M., Schuler G.D.,			
RA	Altschul S.F., Zeeberg B., Buertow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatcenco L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Uedlin T.B., Toshiyuki S., Carninci P., Prange C.,			

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Wooley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Vallian D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fillion J., Helton E., Ketterman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalka U., Smillius D.E.,  
 RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.",  
 RT Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RL [6]  
 RN SEQUENCE OF 28-424 FROM N.A.  
 RP TISSUE=Blood;  
 RC MEDLINE=90049640; PubMed=1425921;  
 RX Fomesgaard A., Hirsch V.M., Johnson P.R.,  
 RA "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.",  
 RT Eur. J. Immunol. 22:2973-2981(1992).  
 RL [7]  
 RN SEQUENCE OF 26-394 AND DISULFIDE BOND.  
 RP MEDLINE=90078232; PubMed=2592374;  
 RX Carr S.A., Hemling M.E., Folena-Waserman G., Sweet R.W., Anumula K.,  
 RA Barr J.R., Huddleston M.J., Taylor P.,  
 RT "Protein and carbohydrate structural analysis of a recombinant soluble  
 RT CD4 receptor by mass spectrometry.",  
 RL J. Biol. Chem. 264:21286-21295(1989).  
 RN [8]  
 RP SEQUENCE OF 26-40.  
 RX PubMed=15340161; DOI=10.1110/ps.04682504;  
 RA Zhang Z., Henzel W.J.,  
 RT "Signal peptide prediction based on analysis of experimentally  
 RT verified cleavage sites",  
 RT Protein Sci. 13:2819-2824(2004).  
 RN [9]  
 RP REMOVAL FROM CELL SURFACE BY HIV-1 NEF, AND MUTAGENESIS OF MET-432;  
 RP SER-433; 438-LEU-LEU-439 AND SER-440.  
 RX PubMed=8124721;  
 RA Aiken C., Konner J., Landau N.R., Lenburg M.E., Trono D.,  
 RT "Nef induces CD4 endocytosis: requirement for a critical dileucine  
 RT motif in the membrane-proximal CD4 cytoplasmic domain.",  
 RL Cell 76:853-864(1994).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.  
 RX MEDLINE=1061881; PubMed=1701030; DOI=10.1038/348411a0;  
 RA Wang J., Yan Y., Garret T.P., Liu J., Rodgers D.W., Garlick R.L.,  
 RA Tarr G.E., Husain Y., Reinherz E.L., Harrison S.C.,  
 RT "Atomic structure of a fragment of human CD4 containing two  
 RT immunoglobulin-like domains",  
 RL Nature 348:411-418(1990).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.  
 RX MEDLINE=1061882; PubMed=2247146; DOI=10.1038/348419a0;  
 RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Atchos J.,  
 RA Rosenberg M., Dai X., Xiong N.-H., Axel R., Sweet R.W.,  
 RA Hendrickson W.A.,  
 RT "Crystal structure of an HIV-binding recombinant fragment of human  
 RT CD4.",  
 RL Nature 348:419-426(1990).  
 RN [12]  
 RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.  
 RX MEDLINE=97311402; PubMed=9168119;  
 RA Wu H., Kwong P.D., Hendrickson W.A.,  
 RT "Dimeric association and segmental variability in the structure of  
 RT human CD4.",  
 RL Nature 387:527-530(1997).  
 RN [13]  
 RP PALMITOYLATION.  
 RX MEDLINE=92317088; PubMed=1618861;  
 RA Crise B., Rose U.K.,

RT "Identification of palmitoylation sites on CD4, the human  
 RT immunodeficiency virus receptor.",  
 RL J. Biol. Chem. 267:13593-13597(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein. Removed from plasma  
 CC membrane by HIV-1 Nef protein that increases clathrin-dependent  
 CC endocytosis of this antigen to target it to lysosomal degradation.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -1- DATABASE: NAME=PROIM; NOTE=CD guide CD4 entry;  
 CC WWW="http://www.ncbi.nlm.nih.gov/prov/cd/cd4.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by, and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M12807; AAA35572.1; -  
 DR EMBL: U47924; AAB51309.1; -  
 DR EMBL: M35160; AAL16068.1; -  
 DR EMBL: BC025782; AAH25782.1; -  
 DR PIR: A90872; RWHUT4.  
 DR PDB: 1CDH; X-ray; @=26-203.  
 DR PDB: 1CDI; X-ray; @=25-203.  
 DR PDB: 1CDJ; X-ray; @=26-203.  
 DR PDB: 1CDU; X-ray; @=26-203.  
 DR PDB: 1CDV; X-ray; @=26-203.  
 DR PDB: 1G9M; X-ray; C=26-210.  
 DR PDB: 1G9N; X-ray; C=26-210.  
 DR PDB: 1GCI; X-ray; C=26-210.  
 DR PDB: 1UL4; X-ray; D=26-203.  
 DR PDB: 1Q68; NMR; A=421-458.  
 DR PDB: 1WER; NMR; @=427-445.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 1WIP; X-ray; A/B=26-388.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 3CD4; X-ray; @=26-207.  
 DR GlycoSuiteDB; P01730; -  
 DR GeneW; HGNC:1678; CD4.  
 DR H-InvDB; HIX0023001; -  
 DR MIM; 186940; -  
 DR GO: GO:0005886; C:Plasma membrane; TAS.  
 DR GO: GO:0042101; C:T-cell receptor complex; NAS.  
 DR GO: GO:0015026; F:coreceptor activity; NAS.  
 DR GO: GO:0015029; F:internalization receptor activity; TAS.  
 DR GO: GO:0042289; F:MHC class II protein binding; NAS.  
 DR GO: GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO: GO:0006955; P:immune response; NAS.  
 DR GO: GO:0045085; P:positive regulation of interleukin-2 biosyn. . .; NAS.  
 DR GO: GO:0030217; P:T-cell differentiation; NAS.  
 DR GO: GO:0045058; P:T-cell selection; NAS.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; NAS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR000973; CD4 TCAG.  
 DR InterPro: IPR007110; Ig-like.  
 DR InterPro: IPR003596; Ig-V.  
 DR Pfam: PF00047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV\_1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW 3D-structure; Direct protein sequencing; Glycoprotein;  
 KW Immune response; Immunoglobulin domain; Lipoprotein; Palmitate;  
 KW Polymorphism; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458 T-cell surface glycoprotein CD4.  
 FT DOMAIN 26 396 Extracellular (Potential).  
 FT TRANSMEM 397 418 Potential.  
 FT DOMAIN 419 458 Cytoplasmic (Potential).

```

FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.
Query Match 39.9%; Score 1064; DB 1; Length 458;
Best Local Similarity 100.0%; Pred. No. 1.8e-59;
Matches 204; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPAATQGGKVVGLKGGDTVELCTASQKSIQPHMKNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPAATQGGKVVGLKGGDTVELCTASQKSIQPHMKNQIK 60
QY 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKMLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKMLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
QY 181 TWCTVLOKQKVEFKIDIVVLAFOKAS 208
DB 181 TWCTVLOKQKVEFKIDIVVLAFOKAS 208

RESULT 2
CD4_PANTR STANDARD; PRT: 458 AA.
AC P16004;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Pan troglodytes (Chimpanzee).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Pan.
OX NCBI_TaxId=9598;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=30182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Cameron D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site."
RT Cell 60:747-754(1990).
RN [2]
RP SEQUENCE OF 26-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=3049640; PubMed=1425921;
RA Fomsgard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus."
RT Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements
CC or send an email to license@ebi.ac.uk).
CC -----
DR EMBL; M3135; AAA35407.1; -
DR EMBL; X73323; CAAS1749.1; -
DR PIR; B32722; RWC274.
DR HSSP; P01730; IWIO.

```

```

DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006555; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR006424; CD2.
DR InterPro; IPR006973; CD4_TCSAg.
DR InterPro; IPR007110; Ig_V.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG_LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSMEM 397 418
FT DOMAIN 419 458
FT DOMAIN 126 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 42 42
FT CONFLICT 62 62
FT CONFLICT 191 191
SQ SEQUENCE 458 AA; 51057 MW; A7C3AC8A5257D3AD CRC64;

Query Match 39.0%; Score 1040; DB 1; Length 458;
Best Local Similarity 98.1%; Pred. No. 5.9e-58;
Matches 204; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALPPAATQGGKVVGLKGGDTVELCTASQKSIQPHMKNQIK 60
DB 1 MNRGVPFRHLVLVQLALPPAATQGGKVVGLKGGDTVELCTASQKSIQPHMKNQIK 60
QY 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKMLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILNGQSFLLTKGSPSKLNDRADSRSLMDQGNFPLIIKMLKIEDSDTYICEVEDQKEEVOL 120
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRKNIOGKTLVSQLELDG 180
QY 181 TWCTVLOKQKVEFKIDIVVLAFOKAS 208
DB 181 TWCTVLOKQKVEFKIDIVVLAFOKAS 208

RESULT 3
CD4_MACPA STANDARD; PRT: 458 AA.
AC P79185;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

```

CC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9541;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymocytes;  
 RA Tatsumi M., Yabe M., Yamada Y.K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63349; BAA09673.1; -.  
 DR HSSP; P01730; 1MBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TCAG.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF0047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KM Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT STGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 203 203  
 FT DOMAIN 317 317  
 FT DOMAIN 374 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 SQ SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;

Query Match 35.74; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.94; Pred. No. 1.9e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 121 LVFGLTANSDTHLLOQSLLTLLESPGSSPSVQCRSPGRKNIQGGKTLVSQSLQDSG 180  
 DB 121 LVFGLTANSDTHLLEQSLLTLLESPGSSPSVKRSPGKNIQGGRTLSVPQLERDSG 180  
 QY 181 TWTCYTLQNKQKVEPKIDIVLAFQKAS 208  
 DB 181 TWTCVSDQKVEFKIDIVLAFQKAS 208  
 RESULT 4  
 CD4\_MACFU STANDARD; PRT; 458 AA.  
 AC P79184;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
 DE T/Leu-3).  
 GN Name=CD4;  
 OS Macaca fuscata fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;  
 OC Cercopithecinae; Macaca.  
 OX NCBI\_TaxID=9543;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hashimoto O., Tatsumi M.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63348; BAA09672.1; -.  
 DR HSSP; P01730; 1MBR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TCAG.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF0047; Ig\_2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KM Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT STGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 203 203  
 FT DOMAIN 317 317  
 FT DOMAIN 374 374  
 FT CARBOHYD 42 42

```

FT CARBOHYD 296 296 N-linked (GlcNAc... ) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc... ) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7EF08185535 CRC64;

Query Match 35.6%; Score 951; DB 1; Length 458;
Best Local Similarity 86.5%; Pred. No. 2,66-52;
Matches 184; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVLQALALPAATGKRVYLGKGGDYELTCTASOKKSIOFHKNSNOIK 60
D 1 MNRGIPFRHLIVLQALALPAATGKRVYLGKGGDYELTCTASOKKSIOFHKNSNOIK 60
QY 61 ILNGSGFLTKGSPKSLNDRADSRRLMDQGNPLIINKLKIEDSDTYICEVEDQKEEVL 120
D 61 ILGIQSGFLTKGSPKSLNDRADSRRLMDQGNPLIINKLKIEDSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSTHLLQSGSLTTLTSSPPGSSPVCCRPGRKNIOGKTLISVQLELDDSG 180
D 121 LVFGLTANSTHLLQSGSLTTLTSSPPGSSPVCCRPGRKNIOGKTLISVQLELDDSG 180
QY 181 TWICTVYLNQKQVFKIDIVYLAFOKAS 208
D 181 TWICTVYLNQKQVFKIDIVYLAFOKAS 208
D 181 TWICTVYLNQKQVFKIDIVYLAFOKAS 208

RESULT 5
CD4 - MACMU STANDARD; PRT; 458 AA.
ID CD4 - MACMU STANDARD; PRT; 458 AA.
AC P16T03; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca mulatta (Rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxID=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024; DOI=10.1016/0092-8674(90)90089-W;
RA Camerini D., Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Thymocytes;
RA Hashimoto O., Tatum M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 22:2973-2981(1992).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.B., Disotell T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";

```

```

RL Mol. Biol. Evol. 15:892-900(1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL, with the
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; M31134; AAA36838.1; -.
DR EMBL; D63347; BAA09671.1; -.
DR EMBL; X73326; CAA51752.1; -.
DR EMBL; AF057385; AAC25129.1; -.
DR HSP; P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0066955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR000973; CD4_TcAg.
DR InterPro; IPR007110; Ig-Like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_Like; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458 T-cell surface glycoprotein CD4.
FT DOMAIN 26 396 Extracellular (Potential).
FT TRANSMEM 397 418 Potential.
FT DOMAIN 419 458 Cytoplasmic (Potential).
FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.
FT DOMAIN 204 317 Ig-like C2-type 2.
FT DOMAIN 318 374 Ig-like C2-type 3.
FT CARBOHYD 296 296 N-linked (GlcNAc... ) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc... ) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 328 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 42 N -> T (in Ref. 1).
FT CONFLICT 62 62 L -> S (in Ref. 3).
FT CONFLICT 67 67 L -> S (in Ref. 2).
FT CONFLICT 169 169 I -> L (in Ref. 2).
FT CONFLICT 191 191 K -> N (in Ref. 2).
FT CONFLICT 248 248 S -> P (in Ref. 2).
FT CONFLICT 265 265 R -> Q (in Ref. 3).
FT CONFLICT 349 349 A -> T (in Ref. 2).
SQ SEQUENCE 458 AA; 50884 MW; 8BB80339FAFEC08 CRC64;

Query Match 35.4%; Score 945; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 6,2e-52;
Matches 183; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLIVLQALALPAATGKRVYLGKGGDYELTCTASOKKSIOFHKNSNOIK 60
D 1 MNRGIPFRHLIVLQALALPAATGKRVYLGKGGDYELTCTASOKKSIOFHKNSNOIK 60

```

```

QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGIQGFHLTKGPKSLNDRADSRSLMDQGSFMIINKLKIEDSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSDTHLLEQOSLTLTLESPGSSPSVQCSPPGKNIQGGKTLVSQLEODSG 180
DB 121 LVFGLTANSDTHLLEQOSLTLTLESPGSSPSVQCSPPGKNIQGGKTLVSQLEODSG 180
QY 181 TWTCVLOQOKVEFKIDIVLAFOKAS 208
DB 181 TWTCVLOQOKVEFKIDIVLAFOKAS 208

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
ID 008340; P79196;
AC 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (Pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsuami M.;
RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RC TISSUE=Blood;
RX MEDLINE=93049640; PubMed=1425921;
RA Fomsgard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Eur. J. Immunol. 23:2973-2981(1992).
CC -FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -SUBUNIT: Associates with p56-lck (by similarity).
CC -SUBCELLULAR LOCATION: Type I membrane protein.
CC -SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL; D63346; BAA09670.1; -
CC EMBL; X73325; CAA51751.1; -
CC HSSP; P01730; IWR.
CC GO; GO:0042101; C:T-cell receptor complex; ISS.
CC GO; GO:0015026; F:coreceptor activity; ISS.
CC GO; GO:0042289; F:MHC class II protein binding; ISS.
CC GO; GO:0006955; P:immune response; ISS.
CC GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
CC GO; GO:0030217; P:T-cell differentiation; ISS.
CC GO; GO:0045058; P:T-cell selection; ISS.
CC GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
CC InterPro; IPR008424; CD2.
CC InterPro; IPR000973; CD4_TcRA.
CC InterPro; IPR007110; Ig-like.
CC InterPro; IPR003596; Ig_v.
CC Pfam; PF00047; Ig_2.
CC PRINTS; PR00692; CD4TCANTIGEN.

```

```

DR SMART; SMO0406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
KM Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KM Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT STGNL 1 25
FT CHAIN 26 458
FT TRANSMEM 26 396
FT DOMAIN 397 418
FT DOMAIN 419 458
FT DOMAIN 26 125
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 57 57
FT CONFLICT 91 91
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2C8B3EB16 CRC64;

Query Match 35.4%; Score 944; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 7, 2e-52;
Matches 183; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALPAATQGRKVVLGKKGVETVETCTASQKKSIOFHWKNSNQIK 60
DB 1 MNRGVPFRHLVLTQALPAATQGRKVVLGKKGVETVETCTASQKKSIOFHWKNSNQIK 60
QY 61 ILGNQGSFLTKGPKSLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120
DB 61 ILGIQGFHLTKGPKSLNDRADSRSLMDQGSFMIINKLKIEDSDTYICEVENKEEVEL 120
QY 121 LVFGLTANSDTHLLEQOSLTLTLESPGSSPSVQCSPPGKNIQGGKTLVSQLEODSG 180
DB 121 LVFGLTANSDTHLLEQOSLTLTLESPGSSPSVQCSPPGKNIQGGKTLVSQLEODSG 180
QY 181 TWTCVLOQOKVEFKIDIVLAFOKAS 208
DB 181 TWTCVLOQOKVEFKIDIVLAFOKAS 208

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
ID 008338; O07593; Q28217;
AC 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Griivet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsuami M.;
RL "Molecular cloning and expression of african green monkey CD4.";
RT Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RC TISSUE=Blood;

```

RX MEDLINE=93049640; Pubmed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [3]  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RX MEDLINE=98017879; Pubmed=9379478;  
 RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT Relation between phylogeny of African green monkey CD4 genes and  
 RT their respective simian immunodeficiency virus genes";  
 RL J. Med. Primatol. 26:120-128(1997).  
 RN [4]  
 RP SEQUENCE OF 107-192 FROM N.A.  
 RX MEDLINE=98320644; Pubmed=9656488;  
 RA Harris E.E., Disotell T.R.;  
 RT "Nuclear gene trees and the phylogenetic relationships of the  
 RT mangabeys (Primates: Papionini).";  
 RL Mol. Biol. Evol. 15:892-900(1998).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D86589; BAAL132.1; -;  
 DR EMBL; X73322; CAAS1748.1; -;  
 DR EMBL; AF001226; AAB60873.1; -;  
 DR EMBL; AF001228; AAB60875.1; -;  
 DR EMBL; AF057380; AAC25124.1; -;  
 DR HSSP; P01710; IWIQ.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; IG\_1Ike.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT CHAIN 1 25  
 FT DOMAIN 26 458  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 281 281  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109

FT DISULFID 155 184 By similarity.  
 FT DISULFID 328 370 By similarity.  
 FT LIPID 419 439 S-palmitoyl cysteine (By similarity).  
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).  
 FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).  
 FT CONFLICT 59 46 I -> T (in Ref. 3; AAB60873).  
 FT CONFLICT 115 115 K -> E (in Ref. 1).  
 FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).  
 FT CONFLICT 200 200 M -> L (in Ref. 2 and 3).  
 FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).  
 FT CONFLICT 271 271 K -> E (in Ref. 3; AAB60873).  
 FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).  
 SQ SEQUENCE 458 AA; 51158 MW; FCS23D2EDD1F72E7 CRC64;  
 Query Match 34.7%; Score 926; DB 1; Length 458;  
 Best Local Similarity 86.5%; Pred. No. 1e-50;  
 Matches 180; Conservative 12; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLVLQALPAATQGGKVVILGKKGVDELCTASQKKSIQFMKNSNQIK 60  
 DB 1 MMNGIFRHLVLQALPAATQGGKVVILGKKGVDELCTASQKKSIQFMKNSNQIK 60  
 QY 61 ILGNQSFITKGPSSKLNDRPADSRRLMDQGNPFLIKLIKEDSDTITCEVEDQKEVOL 120  
 DB 61 ILGKQSFITKGPSSKLNDRPADSRRLMDQGNPFLIKLIKEDSDTITCEVEDQKEVOL 120  
 QY 121 LVFGLRNSDTHLQGGSLTLTSPSPSSPSQCRSPRGKNIQGGKTLVSQLELDPSG 180  
 DB 121 LVFGLRNSDTHLQGGSLTLTSPSPSSPSQCRSPRGKNIQGGKTLVSQLELDPSG 180  
 QY 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCVTLQNKQKVEFKIDIVLAFQKAS 208  
 RESULT 8  
 Q652C8 PRELIMINARY; PRT; 244 AA.  
 ID 0652C8;  
 AC 25-OCT-2004 (T-EMBLrel. 28, Created)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (T-EMBLrel. 28, Last annotation update)  
 DE Single-chain Fv (Fragment).  
 GN Name=scFv;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97362799; Pubmed=9219263;  
 RA Kontermann R.E., Wing M.G., Winter G.;  
 RT "Complement recruitment using bispecific diabodies.";  
 RL Nat. Biotechnol. 15:629-631(1997).  
 DR EMBL; Y13057; CAA73500.1; -;  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig\_1Ike.  
 DR InterPro; IPR003596; Ig\_V.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00409; IGV; 2.  
 DR SMART; SM00406; IGV; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 FT NON\_TER 1 244  
 FT NON\_TER 244 244  
 SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;  
 Query Match 30.9%; Score 824.5; DB 2; Length 244;  
 Best Local Similarity 64.7%; Pred. No. 1.3e-44;  
 Matches 163; Conservative 29; Mismatches 51; Indels 9; Gaps 2;  
 QY 244 QVQLLESGAEVKKPGSSSVYSCASGDTFIRYSFTWVROAPQGLBWKGRITITLIDVAHY 303  
 DB 1 QVQLVDSGAEVKKPGSSSVYSCASGDTFIRYSFTWVROAPQGLBWKGRITITLIDVAHY 303



```

RT Immunodeficiency virus."
RL Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.ebi.ac.uk/announcements/
CC or send an email to license@ebi.ac.uk).
CC -----
CC EMBL: X73324; CAAS1750.1; -.
DR HSSP; P01730; 1WIO.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006935; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR GlycoProtein; Immune response; Immunoglobulin domain; Lipoprotein;
DR Palmitate; Repeat; T-cell; Transmembrane.
DR NON_TER 1
DR DOMAIN <1 369 Extracellular (Potential).
DR TRANSMEM 370 391 Potential.
DR DOMAIN 392 >397 Cytoplasmic (Potential).
DR DOMAIN <1 98 Ig-like V-type.
DR DOMAIN 99 176 Ig-like C2-type 1.
DR DOMAIN 177 290 Ig-like C2-type 2.
DR DOMAIN 291 347 Ig-like C2-type 3.
DR CARBOHYD 269 298 N-linked (GlcNAc. . .) (By similarity).
DR DISULFID 14 82 N-linked (GlcNAc. . .) (By similarity).
DR DISULFID 128 157 By similarity.
DR DISULFID 301 343 By similarity.
DR LIPID 392 393 S-palmitoyl cysteine (By similarity).
DR LIPID 395 395 S-palmitoyl cysteine (By similarity).
DR NON_TER 397 397
DR SEQUENCE 397 AA; 44081 MW; 67887397A6B7EA4F CRC64;

Query Match 30.7%; Score 819; DB 1; Length 397;
Best Local Similarity 87.3%; Pred. No. 5.2e-44;
Matches 156; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

```

```

DB 181 S 181
RESULT 11
ID 009261 PRELIMINARY; PRT; 397 AA.
AC 009261;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE CD4 (Fragment).
OS Cercopithecus sabaues.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopitheciinae; Cercopithecus.
OC NCBI_TaxID=60711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes."
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL; AF001225; AAB60872.1; -.
DR HSSP; P01730; 1WIO.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0007155; P:cell adhesion; IEA.
DR GO; GO:0006955; P:immune response; IEA.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR00973; CD4_TcRg.
DR InterPro; IPR007110; Ig-like.
DR Pfam; PF05790; C2-setc; 2.
DR Pfam; PF00047; Ig; 1.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PS50835; IG_LIKE; 1.
DR NON_TER 1
DR NON_TER 397 397
DR SEQUENCE 397 AA; 43945 MW; 80C963B92A68CD3 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;
Best Local Similarity 86.2%; Pred. No. 5.3e-43;
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

```

OS Cercopithecus tantalus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopithecinae; Cercopithecus.  
 NCBI\_Taxid=60712;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98017879; PubMed=9379478;  
 RA Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001221; AAB60868.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IgV; 1.  
 DR PROSITE; PS50835; Ig\_LIKE; 1.  
 FT NON\_TER 1 397  
 FT TER 397  
 SQ SEQUENCE 397 AA; 43954 MW; CF7F2F5D82335BD0 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKKDVTVELTCTASOKSIOFHWKNSNOIKILGQGSFLTKGSKLNDRAISRSLW 87  
 DB 1 VVLGKKDVTVELTCNASQNTTTFHWKNSNOIKILGQGSFLTKGSKLNDRAISRSLW 60  
 QY 88 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLQGSILTLTLESPP 147  
 DB 61 DQCFSMIILKNLKIETSEYICEVENKEEVELVFGLTANSPTHLQGSILTLTLESPP 120  
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLAQOKVEFKIDIVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLAQOKVEFKIDIVLAFOKA 180  
 QY 208 S 208  
 DB 181 S 181

RESULT 13  
 009263 PRELIMINARY; PRT; 397 AA.  
 AC 009263;  
 DT 01-JUL-1997 (TREMBLrel. 04, Created)  
 DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE CD4 (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus tantalus.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopithecinae; Cercopithecus.  
 NCBI\_Taxid=60712;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98017879; PubMed=9379478;  
 RA Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";

RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001222; AAB60869.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IgV; 1.  
 DR PROSITE; PS50835; Ig\_LIKE; 1.  
 FT NON\_TER 1 397  
 FT TER 397  
 SQ SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKKDVTVELTCTASOKSIOFHWKNSNOIKILGQGSFLTKGSKLNDRAISRSLW 87  
 DB 1 VVLGKKDVTVELTCNASQNTTTFHWKNSNOIKILGQGSFLTKGSKLNDRAISRSLW 60  
 QY 88 DQGNFPLIKNLKIEDSDTYICEVEDQKEEVLVFGLTANSPTHLQGSILTLTLESPP 147  
 DB 61 DQCFSMIILKNLKIETSEYICEVENKEEVELVFGLTANSPTHLQGSILTLTLESPP 120  
 QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLAQOKVEFKIDIVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLAQOKVEFKIDIVLAFOKA 180  
 QY 208 S 208  
 DB 181 S 181

RESULT 14  
 095NE9 PRELIMINARY; PRT; 397 AA.  
 AC 095NE9;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE CD4 (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus pygerythrus (Vervet monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;  
 OC Cercopithecinae; Cercopithecus.  
 NCBI\_Taxid=60710;  
 RN  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=98017879; PubMed=9379478;  
 RA Fomesgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 DR EMBL; AF001227; AAB60874.1; -.  
 DR HSSP; P01730; 1WIO.  
 DR GO; GO:0016021; C:integral to membrane; IEA.  
 DR GO; GO:0007155; P:cell adhesion; IEA.  
 DR GO; GO:0006955; P:immune response; IEA.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig v.  
 DR Pfam; PF005790; C2-sec; 2.  
 DR Pfam; PF00047; Ig; 1.

DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG\_LIKE; 1.  
 FT NON\_TER 1  
 FT SEQUENCE 397 AA; 43946 MW; 21C3B308B2AFBEC0 CRC64;

Query Match 30.1%; Score 803; DB 2; Length 397;  
 Best Local Similarity 86.2%; Pred. No. 5.3e-43;  
 Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNGSFLTKGFSKLANDRADSRSLW 87  
 DB 1 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNGSFLTKGFSKLANDRADSRSLW 60

QY 88 DQGNFPLITKNLKIETSDTYICEVEDQKEVQLVFGLTNSDTHLLOQSFLTLESPP 147  
 DB 61 DQGNFPLITKNLKIETSDTYICEVEDQKEVQLVFGLTNSDTHLLOQSFLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNGKVEFKIDIVVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNGKVEFKIDIVVLAFOKA 180

QY 208 S 208  
 DB 181 S 181

## RESULT 15

009259 ID 009259 PRELIMINARY; PRT; 397 AA.

DT 01-JUL-1997 (TREMBlrel. 04, Created)

DT 01-JUL-1997 (TREMBlrel. 04, Last sequence update)

DT 01-OCT-2003 (TREMBlrel. 25, Last annotation update)

DE CD4 (Fragment).

GN Name=CD4;

OS Cercopithecus sabaus.

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;

OC Cercopithecinae; Cercopithecus.

OX NCBI\_TaxID=60711;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=98017879; PubMed=9379478;

RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,

Corbet S., Barre-Sinoussi F., Allan J.S.;

RT "Relation between phylogeny of African green monkey CD4 genes and

their respective simian immunodeficiency virus genes.";

RL J. Med. Primatol. 26:120-128(1997).

DR EMBL; AF001223; AAB60870.1; -.

DR HSSP; P01730; IMIQ.

DR GO; GO:0016021; C:integral to membrane; IEA.

DR GO; GO:0007155; P:cell adhesion; IEA.

DR GO; GO:0006955; P:immune response; IEA.

DR InterPro; IPR008424; CD2.

DR InterPro; IPR00973; CD4\_TAg.

DR InterPro; IPR007110; IG\_Like.

DR Pfam; PF05790; C2-set; 2.

DR PRINTS; PR00692; CD4TCANTIGEN.

DR SMART; SM00406; IGV; 1.

DR PROSITE; PSS0835; IG\_LIKE; 1.

FT NON\_TER 1

FT SEQUENCE 397 AA; 43881 MW; 7CB39AD0F8506C81 CRC64;

QY 29.9%; Score 799; DB 2; Length 397;  
 Best Local Similarity 85.6%; Pred. No. 9.6e-43;  
 Matches 155; Conservative 11; Mismatches 15; Indels 0; Gaps 0;

QY 28 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNGSFLTKGFSKLANDRADSRSLW 87

DB 1 VVLGKGDVLTCTASQKSIOPHWKNSNOIKILNGSFLTKGFSKLANDRADSRSLW 60

QY 88 DQGNFPLITKNLKIETSDTYICEVEDQKEVQLVFGLTNSDTHLLOQSFLTLESPP 147  
 DB 61 DQGNFPLITKNLKIETSDTYICEVEDQKEVQLVFGLTNSDTHLLOQSFLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNGKVEFKIDIVVLAFOKA 207  
 DB 121 GSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGTWCTVLQNGKVEFKIDIVVLAFOKA 180

QY 208 S 208  
 DB 181 S 181

Search completed: September 20, 2005, 16:09:30  
 Job time : 141 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 08:29:41 / Search time 6457 Seconds  
(without alignments)  
10806.181 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440  
Sequence: 1 atgaaccgsggagtcctt.....gtcagcagtaacataactgg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database : GenBml:.\*  
1: gb\_ba:.\*  
2: gb\_hcg:.\*  
3: gb\_in:.\*  
4: gb\_om:.\*  
5: gb\_ov:.\*  
6: gb\_pat:.\*  
7: gb\_ph:.\*  
8: gb\_pl:.\*  
9: gb\_pr:.\*  
10: gb\_ro:.\*  
11: gb\_sts:.\*  
12: gb\_sy:.\*  
13: gb\_un:.\*  
14: gb\_vl:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1440	100.0	1440	6	BD268881 Novel ch
2	625	43.4	1213	6	AR363716 Sequence
3	624.4	43.4	1377	6	AX100880 Sequence
4	624.4	43.4	1742	6	AR380468 Sequence
5	624.4	43.4	1742	6	AX287109 Sequence
6	624.4	43.4	1742	9	HUMATCT4
7	624.4	43.4	1910	6	HUMATCT4
8	624.4	43.4	1910	6	108116 Sequence 3
9	624.4	43.4	3133	6	108115 Sequence 1
10	624.4	43.4	7533	12	AY438650 Cloning v
11	622.8	43.2	1273	6	106223 Sequence 1
12	622.8	43.2	1273	6	107147 Sequence 1
13	622.8	43.2	1304	6	AR067943 Sequence
14	622.8	43.2	1304	6	AR166802 Sequence
15	622.8	43.2	1304	6	AR560088 Sequence
16	622.8	43.2	1389	6	AR062468 Sequence
17	622.8	43.2	1389	6	AR067924 Sequence
18	622.8	43.2	1389	6	AR166783 Sequence
19	622.8	43.2	1389	6	AR560069 Sequence

20	622.8	43.2	1537	6	CQ722428 Sequence
21	622.8	43.2	1599	6	AR062469 Sequence
22	622.8	43.2	1599	6	AR067925 Sequence
23	622.8	43.2	1599	6	AR166784 Sequence
24	622.8	43.2	1599	6	AR560070 Sequence
25	622.8	43.2	1728	6	AR062467 Sequence
26	622.8	43.2	1728	6	AR067923 Sequence
27	622.8	43.2	1728	6	AR166782 Sequence
28	622.8	43.2	1728	6	AR560068 Sequence
29	622.8	43.2	1742	6	AR104143 Sequence
30	622.8	43.2	1742	6	108673 Sequence 2
31	622.8	43.2	1742	6	AR339862 Sequence
32	622.8	43.2	1742	6	AR448985 Sequence
33	622.8	43.2	2589	6	AR392157 Sequence
34	622.8	43.2	3084	9	BC025782 Homo sapi
35	621.2	43.1	1273	6	AR104142 Sequence
36	621.2	43.1	1273	6	AR448979 Sequence
37	621.2	43.1	1416	6	107030 Sequence 11
38	621.2	43.1	1416	6	109301 Sequence 1
39	621.2	43.1	1416	6	127613 Sequence 2
40	621.2	43.1	1416	6	AR489592 Sequence
41	621.2	43.1	1742	6	107208 Sequence 4
42	621.2	43.1	2940	6	AR064430 Sequence
43	621.2	43.1	3064	6	109237 Sequence 1
44	621.2	43.1	3426	6	AR064429 Sequence
45	621.2	43.1	3480	6	AR064431 Sequence

#### ALIGNMENTS

RESULT 1  
BD268881 1440 bp DNA linear PAT 17-JUL-2003  
LOCUS Novel chimeric protein for prevention and treatment of HIV  
DEFINITION

ACCESSION BD268881  
VERSION BD268881.1 GI:33078649  
KEYWORDS JP 2002538814-A/1.  
SOURCE synthetic construct  
ORGANISM synthetic construct  
REFERENCE 1 (bases 1 to 1440)  
AUTHORS Berger, E.A. and Castillo, C.D.  
TITLE Novel chimeric protein for prevention and treatment of HIV

#### JOURNAL

Patent: JP 2002538814-A 1 19-NOV-2002;  
THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY  
THE AUSTRIAN NORDIC BIOTHERAPEUTICS AKTIENGESellschaft SECRETARY  
DEPARTMENT OF HEALTH AND HUMAN SERVICES THE NATIONAL INSTITUTES OF  
HEALTH

#### COMMENT

OS Artificial Sequence  
PN JP 2002538814-A/1  
PD 19-NOV-2002  
PF 16-MAR-2000 JP 2000605633  
PR 16-MAR-1999 US 60/124681  
PT EDWARD A BERGER, CHRISTIE DEL CASTILLO  
PC C12N15/09,A61K38/00,A61P31/18,C07K5/103,C07K14/00,C07K14/155,  
PC C07K14/73,C07K16/10,C07K19/00,C12N5/10,C12P21/02// (C12P21/02,  
PC C12N15/00,C12N5/00,A61K37/02  
PC C12N15/00,C12N5/00,A61K37/02  
CC Description of Artificial Sequence: CDA-secFv(17b) FH Key  
Location/Qualifiers  
FT source 1..1440  
FT 1..1440 /organism='Artificial Sequence'.  
Location/Qualifiers  
1..1440 /organism='synthetic construct'  
/mol\_type='genomic DNA'  
/db\_xref='taxon:32630'

#### FEATURES

##### source

#### ORIGIN

Query Match 100.0%; Score 1440; DB 6; Length 1440;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 1 ATGACCGGGAGTCCCTTTTAGCACTTGTGCTGCAACTGGGCGTCCCTCCA 60
   |||
DB 1 ATGACCGGGAGTCCCTTTTAGCACTTGTGCTGCAACTGGGCGTCCCTCCA 60
   |||

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGGATACAGTGAATCAGC 120
   |||
DB 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATCAGC 120
   |||

QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCACTGGAATACTCAACCATTAAG 180
   |||
DB 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCACTGGAATACTCAACCATTAAG 180
   |||

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTGCATCCAGGTGAATGATCGGCT 240
   |||
DB 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTGCATCCAGGTGAATGATCGGCT 240
   |||

QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300
   |||
DB 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300
   |||

QY 301 ATGGAAGACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360
   |||
DB 301 ATGGAAGACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360
   |||

QY 361 CTAGTGTGGATTTGACTGCACTGCAACCCACCTGCTTCAAGGGGCGAGGCTGACC 420
   |||
DB 361 CTAGTGTGGATTTGACTGCACTGCAACCCACCTGCTTCAAGGGGCGAGGCTGACC 420
   |||

QY 421 CTGACCTTTGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGATGAGATCCAGGGGT 480
   |||
DB 421 CTGACCTTTGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGATGAGATCCAGGGGT 480
   |||

QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTAGCTGAGCTCCAGATAGTGGC 540
   |||
DB 481 AAAAACAATACAGGGGGGAAAGACCTCTCGGTGTCTAGCTGAGCTCCAGATAGTGGC 540
   |||

QY 541 ACCGAGCATGCACTGCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGCTG 600
   |||
DB 541 ACCGAGCATGCACTGCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATGCTG 600
   |||

QY 601 GTGTAGCTTTCCAGAAAGGCTCCGAGGTGCGCGTGTGAGGGAGGCGGTTGAGCGGGA 660
   |||
DB 601 GTGTAGCTTTCCAGAAAGGCTCCGAGGTGCGCGTGTGAGGGAGGCGGTTGAGCGGGA 660
   |||

QY 661 GGTGATCCGGTGGCGGAGGTGCGGCGGGGTGGAAGCGGGGTGCGGCTCCGAGGCG 720
   |||
DB 661 GGTGATCCGGTGGCGGAGGTGCGGCGGGGTGGAAGCGGGGTGCGGCTCCGAGGCG 720
   |||

QY 721 GGAAGTTCAGAGGTGCACTGCTGAGTCTGGGGCTGAGGTGAAGAACTGGGCTCTCG 780
   |||
DB 721 GGAAGTTCAGAGGTGCACTGCTGAGTCTGGGGCTGAGGTGAAGAACTGGGCTCTCG 780
   |||

QY 781 GTGAGGTCTCTGCAAGGCTCTGAGACACTTTCATCAGATTAATTTAACTGGG 840
   |||
DB 781 GTGAGGTCTCTGCAAGGCTCTGAGACACTTTCATCAGATTAATTTAACTGGG 840
   |||

QY 841 CGACAGGCCCCCTGCAAGGCTTGAAGTGAAGGAGATCATCACTTCTTGAAGTA 900
   |||
DB 841 CGACAGGCCCCCTGCAAGGCTTGAAGTGAAGGAGATCATCACTTCTTGAAGTA 900
   |||

QY 901 GCACATCAGCAACCGCACTCCAGGGGAGATCAAGTTACCGGGCAAGTCCACAGC 960
   |||
DB 901 GCACATCAGCAACCGCACTCCAGGGGAGATCAAGTTACCGGGCAAGTCCACAGC 960
   |||

QY 961 ACAGTCTACCTGAGCTGCGGAATCTAAGATCTGACGATACGCGGCTATATTTCTGTGCG 1020
   |||
DB 961 ACAGTCTACCTGAGCTGCGGAATCTAAGATCTGACGATACGCGGCTATATTTCTGTGCG 1020
   |||

QY 1021 GAGGTGTACGAGGAGGCGGACGAGGGGGAATATGATTAATGAGGTTCTGAAACAT 1080
   |||
  
```

RESULT 2  
 AR363716 1213 bp DNA linear PAT 03-SEP-2003  
 LOCUS AR363716 Sequence 1 from patent US 5223418.  
 DEFINITION AR363716  
 ACCESSION AR363716  
 VERSION AR363716.1 GI:34425655  
 KEYWORDS  
 .  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE 1 (bases 1 to 1213)  
 AUTHORS Arcuri,E.J., Brawner,M.E., Donovan,M.J., Gerber,R.G. and Keller,J.A.  
 TITLE Method of improving the yield of heterologous proteins produced by Streptomyces lividans  
 JOURNAL Patent: US 5223418-A 1 29-JUN-1993;  
 FEATURES  
 1..1213  
 /organism="Unknown"  
 /mol\_type="genomic DNA"

ORIGIN

Query Match 43.4%; Score 625; DB 6; Length 1213;  
 Best Local Similarity 97.7%; Pred. No. 1,3e-134;  
 Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

```

QY 1 ATGACCGGGAGTCCCTTTTAGCACTTGTGCTGCAACTGGGCGTCCCTCCA 60
   |||
DB 79 ATGACCGGGAGTCCCTTTTAGCACTTGTGCTGCTGCAACTGGGCGTCCCTCCA 138
   |||

QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATCAGC 120
   |||
DB 139 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAATCAGC 198
   |||

QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCACTGGAATACTCAACCATTAAG 180
   |||
DB 199 TGTACAGCTTCCAGAAAGAGCATTCATTCCACTGGAATACTCAACCATTAAG 258
   |||

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTGCATCCAGGTGAATGATCGGCT 240
   |||
DB 259 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTGCATCCAGGTGAATGATCGGCT 318
   |||

QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300
   |||
DB 319 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 378
   |||
  
```

QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 Db 379 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 438  
 QY 361 CTAGTGTGGGATTTGACTGCGCACTGTGACCAACCCACTGTCTTACGGGGCAGAGCTGACC 420  
 Db 439 CTAGTGTGGGATTTGACTGCGCACTGTGACCAACCCACTGTCTTACGGGGCAGAGCTGACC 498  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 Db 499 CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 558  
 QY 481 AAAAACAATACAGGGGGGAGAAACCCCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGGC 540  
 Db 559 AAAAACAATACAGGGGGGAGAAACCCCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGGC 618  
 QY 541 ACCTGACATGACATGCTCTTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
 Db 619 ACCTGACATGACATGCTCTTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATCGTG 678  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGGAGGTGGCGTGTAGTGGGAGGCG 649  
 Db 679 GTGCTAGCTTTCCAGAAAGGCTCCGGAGGTGGCGTGTAGTGGGAGGCG 727

## RESULT 3

AX100880

LOCUS AX100880 1377 bp DNA linear PAT 10-APR-2001  
 DEFINITION Sequence 30 from Patent WO0122084.  
 AX100880  
 AX100880.1 GI:13619792

## KEYWORDS

ORGANISM Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euteria; Primates; Catarrhini; Homidae; Homo.

## REFERENCE

1 Jakobsen, B. K.  
 SPr identification of inhibitors of receptor-ligand interactions  
 Patent: WO 0122084-A 30 29-MAR-2001;  
 Avidex Ltd (GB)

## FEATURES

source location/Qualifiers  
 1..1377  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..1377  
 /note="unassigned protein product"  
 /codon\_start=1  
 /protein\_id="CAC36495.1"  
 /db\_xref="GI:13619793"

## CDS

translation="MNRGVPFRLLLVLLALPAATQGRKVVIGKKDTEVETLTAS  
 OKSIOFHMKNISNOIKILNQSFLTGPESKLNDRADSRSLMDQGNPLIKNLKIE  
 DSDTYICEVDQKEVQLVGLTANSDTLQOSLTLLESPGSPVQCSPPG  
 KNLOGKTLVSOLELDSGTMTCTVONKKVEFKDIVLAFKASSTIVYKGEBO  
 VESFPLATPVEKLTSGELMMOABRASSKSWTTPDKKVEYKRVTDOPKLOMK  
 KLPLHLTPQALPOYAGSGNLTALAEKTKLQEVNLVYMRATOLQKNITCEVMGT  
 SPKMLSLKLENKARVKRERKAVWVMPKGMQCLSDSGVLLSNTIKVPLWST  
 PVQPMALIVGVAGVALLFTGLGIFCVRCHRRRROERMSOIKRLISEKTCOCPEH  
 FQTCSP1"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1377;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGTGCAACTGGGCTCTCCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGTGCAACTGGGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGAGTACAGTGAACCTGACC 120  
 Db 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGAGTACAGTGAACCTGACC 120

Db 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGAGTACAGTGAACCTGACC 120  
 QY 121 TGTACAGCTTCCAGAAAGAGATATACATTCCTGAGAAAAAATCTCAACAGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAAAGAGATATACATTCCTGAGAAAAAATCTCAACAGATTAAG 180  
 QY 181 ATTCTGGAAATCAGGCTCTCTTTAACTAAAGTTCATCCAAAGCTGAATGATCGGCT 240  
 Db 181 ATTCTGGAAATCAGGCTCTCTTTAACTAAAGTTCATCCAAAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGAGCTTTGGAGCAAGGAAATTTCCCTGATCATCAAGAAATCTTAAG 300  
 Db 241 GACTCAAGAAAGAGCTTTGGAGCAAGGAAATTTCCCTGATCATCAAGAAATCTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 Db 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGTGTCAATTG 360  
 QY 361 CTAGTGTGGGATTTGACTGCGCACTGTGACCAACCCACTGTCTTACGGGGCAGAGCTGACC 420  
 Db 361 CTAGTGTGGGATTTGACTGCGCACTGTGACCAACCCACTGTCTTACGGGGCAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACGCCCTCATGTGCAATTTAGAGATCCAAAGGGGT 480  
 QY 481 AAAAACAATACAGGGGGGAGAAACCCCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGGC 540  
 Db 481 AAAAACAATACAGGGGGGAGAAACCCCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGGC 540  
 QY 541 ACCTGACATGACATGCTCTTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
 Db 541 ACCTGACATGACATGCTCTTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGGAGGTGGCGTGTAGTGGGAGGCG 649  
 Db 601 GTGCTAGCTTTCCAGAAAGGCTCCGGAGGTGGCGTGTAGTGGGAGGCG 649

## RESULT 4

AR380468

LOCUS AR380468 1742 bp DNA linear PAT 18-DEC-2003  
 DEFINITION Sequence 1013 from patent US 6607879.  
 AR380468  
 AR380468.1 GI:40088102

## KEYWORDS

ORGANISM Unknown.  
 Unknown.  
 Unclassified.

## REFERENCE

1 (bases 1 to 1742)  
 Cocke, B. G., Stuart, S. G. and Seilhamer, J. J.  
 Compositions for the detection of blood cell and immunological  
 response gene expression  
 Patent: US 6607879-A 1013 19-AUG-2003;  
 location/Qualifiers  
 1..1742  
 /organism="Unknown"  
 /mol\_type="genomic DNA"

## FEATURES

source location/Qualifiers  
 1..1742  
 /organism="Unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGTGCAACTGGGCTCTCCCA 60  
 Db 76 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGTGCAACTGGGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGAGTACAGTGAACCTGACC 120  
 Db 136 GCAGCCACTCAGGAGAAAGAGTGTGTGGGCAAAAAGGGAGTACAGTGAACCTGACC 195  
 QY 121 TGTACAGCTTCCAGAAAGAGATATACATTCCTGAGAAAAAATCTCAACAGATTAAG 180

```
Db      196 TGTACAGCTTCCAGAAAGAGCATATTCACACTGAGAAAACTCCAAACGATTAAG 255
Qy      181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCAAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCTCTTTAACTAAAGGTCATCAAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAA 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAA 375
Qy      301 ATGAAGACTCAAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 360
Db      376 ATGAAGACTCAAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 435
Qy      361 CTAGTGTGGATGATGCTGCCTCAACTGTGACCACTGCTTCAGGGGCGAGACCTGACC 420
Db      436 CTAGTGTGGATGATGCTGCCTCAACTGTGACCACTGCTTCAGGGGCGAGACCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGGT 555
Qy      481 AAAAATCAAGAGGGGGGAAAGACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC 540
Db      556 AAAAATCAAGAGGGGGGAAAGACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC 615
Qy      541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db      616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701
```

```
RESULT 5
LOCUS   AX287109      1742 bp      DNA      linear      PAT 21-NOV-2001
DEFINITION
Sequence 7 from Patent WO0164752.
ACCESSION
AX287109
VERSION
AX287109.1 GI:17049085
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
Litman,D.R., Kwon,D., van Kooyk,Y.C. and Geijtenbeek,T.C.
Antibody inhibiting the binding between gp120 and dc-sigin and
screening methods
Patent: WO 0164752-A 7 07-SEP-2001;
JOURNAL
NEW YORK UNIVERSITY (US) ; KATHOLIEKE UNIVERSITEIT NIJMEGEN (NL)
FEATURES
source
1..1742
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
```

```
ORIGIN
Query Match 43.4%; Score 624.4; DB 6; Length 1742;
Best Local Similarity 99.8%; Pred. No. 1.8e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACCGGGAGTCCCTTTAGCACTTGTGTCGCAACTGGGCTCCCTCCA 60
Db 76 ATGAACCGGGAGTCCCTTTAGCACTTGTGTCGCAACTGGGCTCCCTCCA 135
Qy 61 GCAGCCACTCAGGAAAGAAAGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 120
Db 136 GCAGCCACTCAGGAAAGAAAGTGTCTGTGGCAAAAAAGGGATACAGTGAACCTGACC 155
Qy 121 TGTACAGCTTCCAGAAAGAGCATATTCACACTGAGAAAACTCCAAACGATTAAG 180
```

```
Db      196 TGTACAGCTTCCAGAAAGAGCATATTCACACTGAGAAAACTCCAAACGATTAAG 255
Qy      181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGGTCATCCAAAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCTCTTTAACTAAAGGTCATCCAAAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAA 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAA 375
Qy      301 ATGAAGACTCAAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 360
Db      376 ATGAAGACTCAAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 435
Qy      361 CTAGTGTGGATGATGCTGCCTCAACTGTGACCACTGCTTCAGGGGCGAGACCTGACC 420
Db      436 CTAGTGTGGATGATGCTGCCTCAACTGTGACCACTGCTTCAGGGGCGAGACCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCACTGCAATGTAGAGTCCAAAGGGGT 555
Qy      481 AAAAATCAAGAGGGGGGAAAGACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC 540
Db      556 AAAAATCAAGAGGGGGGAAAGACCTCTCCGTGTCTAGCTGAGAGCTCCAGATAGTGGC 615
Qy      541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
Db      616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701
```

```
RESULT 6
LOCUS   HUMATCT4      1742 bp      mRNA      linear      PRI 27-APR-1993
DEFINITION
Human T-cell surface glycoprotein T4 mRNA, complete cds.
ACCESSION
M12807
VERSION
M12807.1 GI:179141
KEYWORDS
glycoprotein; immunoglobulin super gene family; recognition
antigen; surface antigen.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1742)
Maddon,P.J., Litman,D.R., Godfrey,M., Maddon,D.E., Chess,L. and
Axel,R.
The isolation and nucleotide sequence of a cDNA encoding the T cell
surface protein T4; a new member of the immunoglobulin gene family
Cell 42 (1), 93-104 (1985)
JOURNAL
MEDLINE
85254948
PUBMED
2990730
REFERENCE
2 (bases 153 to 153)
Litman,D.R.
Unpublished (1986)
COMMENT
Original source text: Human peripheral T lymphocyte, cDNA to mRNA,
clone pT4B.
T4 is a member of the immunoglobulin supergene family. The T4 mRNA
encodes contiguous V- and J-like elements without the requirement
for DNA recombination events. The V-like domain can be found at
positions 145 to 426 and the J-like domain from 427 to 471. The T4
protein may serve as the specific surface receptor for the AIDS
virus [1].
With the revision of the nucleotide at position 153, the amino acid
encoded by 151-153 becomes lysine. This is now regarded to be the
first residue in the mature protein [2].
Location/Qualifiers
1..1742
/organism="Homo sapiens"
```

FEATURES  
source

```

/mol_type="mRNA"
/db_xref="taxon:9606"
76..1452
/note="T4 surface glycoprotein precursor"
/codon_start=1
/protein_id="AA35572.1"
/db_xref="GI:179142"
/translation="MNRRVPRRHLIVQLALPLAAYGKRVYLAKGKDYIELTCTAS
OKKSIQPHMKNSNQIKILGNQGSFLTKPSKFLNLPADSRSLWPQGNPLITIKULKIE
DSDIYICEVEDQKEBVLVLFGFLANSSTLILDSQSLTSLSPGSSPSVQCSPPGK
KNIGQKFLIVSEIOLSDSGTWCITVLONQKGEKFKIDIVLAQKASIVYKKEGKC
VEFEPPLAFIVETKLTSGSELWMOAERASSKSMITPDLKNEVSKRVTOPKXOMGK
KLPHLITLPOALPOYAGSNLTIALEATGCLHOENVLYNVRATOLQNLTCFVWGP
SPKMLSLKLENKAKRSKRKAVYVNLPEKGMOCCLSDSGVLLSNTIVLPTWST
FVQPMALIVLGVAGLLFLIGLGIFFCVCRHRRROAERMSQIKRLSEKTKCCPPRR
FOKTCSP"
76..150
/note="T4 surface glycoprotein signal peptide"
151..1449
/product="T4 surface glycoprotein mature peptide"
ORIGIN      198 bp upstream of RsaI site.

Query Match      43.4%; Score 624.4; DB 9; Length 1742;
Best Local Similarity 99.8%; Pred. No. 1,8e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 ATGAACCGGGAGTCCCTTTTAGCACTTGTCTGTGCTGCAACTGGCGCTCCTCCA 60
Db      76 ATGAACCGGGAGTCCCTTTTAGCACTTGTCTGTGCTGCAACTGGCGCTCCTCCA 135

QY      61 GCAGCACTCAGGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 120
Db      136 GCAGCACTCAGGGAGAAAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC 195

QY      121 TGTACAGCTTCCAGAAAGACATCAATTCACTGAGAAAACCTCAACAGATAAAG 180
Db      196 TGTACAGCTTCCAGAAAGACATCAATTCACTGAGAAAACCTCAACAGATAAAG 255

QY      181 ATTCTGGAAATCAGAGGCTCCTTCTTAATAAGGTCCTCAAGCTGAATGATCGGCT 240
Db      256 ATTCTGGAAATCAGAGGCTCCTTCTTAATAAGGTCCTCAAGCTGAATGATCGGCT 315

QY      241 GACTCAGAGAAAGCTTTGGGACCAAGAAATTTCCCTGATCATCAGAAATCTTAAG 300
Db      316 GACTCAGAGAAAGCTTTGGGACCAAGAAATTTCCCTGATCATCAGAAATCTTAAG 375

QY      301 ATAGAAGCTCAGATCTTAATCTGTGAATGTGAAGACCAAGAGAGAGGTGCAATTG 360
Db      376 ATAGAAGCTCAGATCTTAATCTGTGAATGTGAAGACCAAGAGAGAGGTGCAATTG 435

QY      361 CTAGTGTTCGGAATTGACTGCCAATCTGACACCACTGCTTCAAGGGGACAGGCTGACC 420
Db      436 CTAGTGTTCGGAATTGACTGCCAATCTGACACCACTGCTTCAAGGGGACAGGCTGACC 495

QY      421 CTGACCTTTGAGAGCCCCCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAAAGGAGT 480
Db      496 CTGACCTTTGAGAGCCCCCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAAAGGAGT 555

QY      481 AAAAAACATACAGGGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTGC 540
Db      556 AAAAAACATACAGGGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGTGC 615

QY      541 ACCTGAAGATGACATGTCTTGACAGACCAAGAAAGAGGTGAGATTCAAAATATGACATGCTG 600
Db      616 ACCTGAAGATGACATGTCTTGACAGACCAAGAAAGAGGTGAGATTCAAAATATGACATGCTG 675

QY      601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

```

LOCUS	HMMATCH4A	1742 bp	DNA	linear	PRI 24-JAN-1998
DEFINITION	Human T4 surface glycoprotein CD4 gene, complete cds.				
ACCESSION	M35160				
VERSION	M35160.1	GI:179143			
KEYWORDS	glycoprotein; immunoglobulin super gene family; recognition antigen; surface antigen.				
SOURCE	Homo sapiens (human)				
ORGANISM	Homo sapiens				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
AUTHORS	1 (bases 1 to 1742)				
TITLE	Hodge, T.W., Sasso, D.R. and McDougal, J.S.				
	Humans with OKT4-epitope deficiency have a single nucleotide base change in the CD4 gene, resulting in substitution of TRP240 for ARG240				
JOURNAL	Hum. Immunol. 30 (2), 99-104 (1991)				
MEDLINE	Hum. 91216786				
PUBMED	1708753				
REFERENCE	2 (bases 1 to 1742)				
AUTHORS	Hodge, T.W.				
JOURNAL	Direct Submission				
	Submitted (12-JUN-1990) T.W. Hodge, A-25 Bldg. 1, Rm. 1226, Center for Disease Control, 1600 Clifton Rd. Atlanta, GA 30333 USA				
COMMENT	Original source text: Human DNA. A cytosine to thymidine transition at nucleotide position 868 results in substitution of TRP-240 for ARG-240.				
FEATURES	source				
	1..1742				
	/organism="Homo sapiens"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:9606"				
	/haplotype="OKT4-negative"				
	76..1452				
	/gene="CD4"				
	76..1452				
	/gene="CD4"				
	/note="T4 surface glycoprotein precursor"				
	/codon_start=1				
	/protein_id="AA16069.1"				
	/db_xref="GI:179144"				
	/translation="MNRGVPFRLHLVLOLALPAAOTGGKVVIGKKDPVELCTAS DSYIOFHMKNNSNOIKILGNQGSFLTGPBKNDRAHSRSIMOGNPILIRIKI BESDTYICEVEDQKEEVOLIVGLFANSDFHLQGSITLLESFPGSSPSVQCRSPRG KNIQGGKTLVSQLELDQSGTWCTYVQNKQVEFKIDIVLAFQKASSIYKKGEGQ VESEFPPLAFVEBKLTSGEELMQAERASSKSWITPDLKNKEVSVKAVTDPKLOMGK KLPLHLTPOLAPVAGSGNLTALAEKTEGLHOEVNLYVMRAITOLQNLCEVWGPT SPKMLSIKLENKEAKVSKREKAVVWVNPAGMOCILSPSGOVLFSNLIKVLPTWST PQGMALIVGVGAAGLLFTGLGTFVYRCRRHRRROAERASQIRKLISSEKTKCCPRP FQKCSPI"				
	76..150				
	/gene="CD4"				
	/note="T4 surface glycoprotein signal peptide"				
	151..1449				
	/gene="CD4"				
	/product="T4 surface glycoprotein mature peptide"				
	792				
	/gene="CD4"				
	/note="g in OKT4 epitope deficiency; c in wt"				
ORIGIN					
Query Match	43.4%; Score 624.4; DB 9; Length 1742;				
Best local Similarity	99.8%; Pred. No. 1.8e-134;				
Matches 625; Conservative	0; Mismatches 1; Indels 0; Gaps 0;				
1	ATGAACCGGGGAGCCCTTTTAGCAATTGCTCTGCTGCAACGCGGCTCTCCCA	60			
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					
26	ATGAACCGGGAGTCCCTTTTAGCACTTCTCTGTGTCGCAACGCGGCTCTCCCA	135			
27					
28					
29					
30					
31					
32					
33					
34					
35					
36					
37					
38					
39					
40					
41					
42					
43					
44					
45					
46					
47					
48					
49					
50					
51	GCAGCCACTCAGGGAAGAAGTGTCTCTGGGCAAAAAGGGGATACAGTGAACCTGACC	120			
52</					

Db 196 TGTACAGCTTCCCAAGAAAGACATTCACATCGGAAAACTCCAAACGAGATAAAG 255  
QY 181 ATCTGGGAAATCAGGGCTCCTTTAACTAAAGGTCCATCCAGCGTAATGATCGGCT 240  
Db 256 ATCTGGGAAATCAGGGCTCCTTTAACTAAAGGTCCATCCAGCGTAATGATCGGCT 315  
QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
Db 316 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 375  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGGTGCATTG 360  
Db 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGGTGCATTG 435  
QY 361 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
Db 436 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 555  
QY 481 AAAAACAATCAGGGGGGAAAGACCTCTCGTGTCTGAGCTGAGACTCCAGATAGTGGC 540  
Db 556 AAAAACAATCAGGGGGGAAAGACCTCTCGTGTCTGAGCTGAGACTCCAGATAGTGGC 615  
QY 541 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGT 600  
Db 616 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGT 675  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 8  
LOCUS 108116 1910 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent EP 0325262.  
ACCESSION 108116  
VERSION 108116.1 GI:589172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1910)  
AUTHORS Brian,S.D.  
TITL Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 3 26-JUL-1989;  
FEATURES  
source 1..1910  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1910;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAATCCCTTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 60  
Db 111 ATGAACCGGGGAATCCCTTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGGATAGTGAAGTGAAC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGGATAGTGAAGTGAAC 230  
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACAGATAAG 180  
Db 231 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACAGATAAG 290  
QY 181 ATCTGGGAAATCAGGGCTCCTTTTAACTAAAGGTCAATCCAGTGAATGATCGGCT 240

Db 291 ATCTGGGAAATCAGGGCTCCTTTTAACTAAAGGTCAATCCAGTGAATGATCGGCT 350  
QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
Db 351 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTG 360  
Db 411 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTG 470  
QY 361 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420  
Db 471 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 530  
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 590  
QY 481 AAAAACAATCAGGGGGGAAAGACCTCTCGTGTCTGAGCTGAGACTCCAGATAGTGGC 540  
Db 591 AAAAACAATCAGGGGGGAAAGACCTCTCGTGTCTGAGCTGAGACTCCAGATAGTGGC 650  
QY 541 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGT 600  
Db 651 ACCTGGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGT 710  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
Db 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 9  
LOCUS 108115 3133 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0325262.  
ACCESSION 108115  
VERSION 108115.1 GI:589171  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 3133)  
AUTHORS Brian,S.D.  
TITL Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 1 26-JUL-1989;  
FEATURES  
source 1..3133  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 3133;  
Best Local Similarity 99.8%; Pred. No. 1.8e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAATCCCTTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 60  
Db 111 ATGAACCGGGGAATCCCTTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGGATAGTGAAGTGAAC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGGATAGTGAAGTGAAC 230  
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACAGATAAG 180  
Db 231 TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATACTCCAAACAGATAAG 290  
QY 181 ATCTGGGAAATCAGGGCTCCTTTTAACTAAAGGTCAATCCAGTGAATGATCGGCT 240  
Db 291 ATCTGGGAAATCAGGGCTCCTTTTAACTAAAGGTCAATCCAGTGAATGATCGGCT 350  
QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

[illegible]

gene

CDS

```

/transl_table=11
/product="gentamicin-3-acetyl transferase"
/protein_id="AAR3802.1"
/db_xref="GI:38197736"
/translation="MLRSSNDVTVQQGSRPRTKLGGSGMGIIRTCRLGDPGVKSRNAL
DLGREGGVDAVTVSOHPDSDYLGNLIRSTFTALAFDDEAVGALAAVLPFEEDP
RSELYIVDLAVSGEHRROGIATLAILNLKHEANALGAYIVIGQVLDGDDPAVALYTKL
GIREVNHFDIPSTAT"
complement(4301..6394)
/gene="CD4-EGFP"
complement(4301..6394)
/gene="CD4-EGFP"
/note="CD4 and enhanced GFP fusion protein"
/codon_start=1
/transl_table=11
/product="CD4-EGFP"
/protein_id="AAR3802.1"
/db_xref="GI:38197737"
/translation="MNRGVPFRHLVLTQLALPAATQGRKVLGGKGDIVELTCTAS
OKSIOPHMNSOKILAGNQSFLTGPGLNDADRSRLSDOGNFPILINKLKE
DPTPIQVEVDOKKEVOLLVFLTANDPTHLGGLSDVITLSESPSPVOCSPRG
KTIQGGKTLTSSQLDELSDSTCTCTYIQLONKQKEFKDIIVLAFQKASLIVYKESQ
VERFSPLAFTVEKTLSSGELMWAERSSKSLITPLKNKESVKTAVTDDPLQJQK
KLPLHLTPALPOLYAGSGNLTALAEKKEKSLQHOEVLVWRATOLQKNTICEWGT
SPTMLSTLKEKAEKAKSKREKAVVWVLPNPAQMOCILSPSQVLESNTKILPTMST
PVPOMALIVLGVAGALLFLGKGFPCVRCRHRRQAEKMSQILRLSEKTCOCPPH
FORTCSIVSKGBELFTGVVPLIVELDEPNVGHKFSVSGEGDAILYGLTLTKICT
TGLTPVWPLTLLTTLTYGQCPSRYDPHMKQHPFKSAMEGVQERTIPKQDGNVK
TRABVREFGDTLVNRIELKGIIDREKDNILGHKLEINYNASHNYIMADKQNGIKNPF
KIRHNIIDGSSVQADHYQONTPIGDGVLLTPDNHYLSTQSALSKDPKEDHVVLEF
VTAGHNLGNDELAYK"

```

# ORIGIN

Query Match	43.48;	Score 624.4;	DB 12;	Length 7533;
-------------	--------	--------------	--------	--------------

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGAGTCCCTTTAGGCACTTGCTTCTGCTGCAACTGGCGCTCTCCA

Db 6394 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTCTGGTGTGCAACTGGGCTCCTCCA 6335

61 GCAGCCACTCAGGGAAGAAGTGGTCTGGGCAAAAAGGGATACAGTGGAACTGACC 120

Db 6334 GCAGCCACTCAGGGAAGTGGTCTGGCAGGGAATACAGTGAAGTACC 6275

QY 121 TGTACAGCTTCCGAGAGAGCATACAATTCCACTGGAAAACTCCAACGAGTAAAG 180

Db 6274 TGTACAGCTTCCGAGAGAGCATACAATTCCACTGGAAAACTCCAACGATTAAG 6215

181 ATTCTGGGAATCAGGCTCCTTCTTAATAAGGTCATCCAAGCTGAATGATCGGCT 240

Db 6214 ATTCTGGGAATCAGGCTCTTCTTAAGGTCCATCCAAGCTGAATGATCGGCT 6155

241 GACTCAGAGAGCGCTTTGGGACCAAGAACTTCCCCCTGATCATCAGAACTTAAG 300

Db 6154 GACTCAAGAGAGCCCTTTGGGACCAAGAACTTCCCCCTGATCATCAAGAACTTAAG 6095

301 ATAGAGACTCAGATACTTACATCTGTGAGTGGAGGACCAGAGGAGGAGGTCATTG 360

Db 6094 ATAGAAGACTCAGATACTTACATCTGTGAGTGGAGGACCAGAAGGAGGAGGTCGAATTG 6035

361 CTAGTGTTCGATTGACTGCCACTCTGACACCCACCTGCTTCAGGGCAGAGCCTGACC 420

Db 6034 CTAGTGTTCGGATGACTGCCAAGCTGACACCCACCTGCTTCAGGGGAGAGCCCTGACC 5975

421 CTGACCTTGAGAGCCCCCTGGTAGTAGCCCCCTCAGTGCATGTAGGAGTCCAAGGGT 480

Db 5974 CTGACCTTGAGAGCCCCCTGGTAGTAGCCCTCAGTGCAATGTAGAGTCCAAAGGGT 5915

481 AAAACATACAGGGGGGAGAACCCTCTCCGTGTCTCAGCTGAGCTCCAGGATAGTGC 540

Db 5914 AAAACATACAGGGGGAACCTCTCCGTGTCTCAGCTGAGCTCCAGGATAGTGC 5855

Oy	541	ACCTGGAATGCACTGTCTTGACAGAAACGAAAGGTGAGCTTCAAAATGACATCGTG	600
Db	5854	ACCTGACATGCACTGTCTTGACAGAACGAAAGGTGAGCTTCAAAATGACATCGTG	5795
Oy	601	GTGCTAGCTTTCAGAAAGGCGCTCCGG	626
Db	5794	GTGCTAGCTTTCAGAAAGGCGCTCCGG	5769
RESULT	11		
LOCUS	106223	1273 bp	DNA
DEFINITION	Sequence 1 from Patent EP 0313377.	linear	PAT 02-DEC-1994
ACCESSION	106223		
VERSION	106223.1	GI:590426	
KEYWORDS	.		
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1273)		
TITLE	Deen, K.C., Folena-Wasserman, G.M., Inacker, R.H. and Sweet, R.W.		
JOURNAL	Process for purification of soluble T4		
FEATURES	Patent: EP 0313377-A1 1 26-Apr-1989;		
source	location/Qualifiers		
	1..1273		

ORIGIN	/mol_type="unassigned DNA"				
Query Match	43.2%	Score 622.8;	DB 6;	Length 1273;	
Best Local Similarity	99.7%;	Pred. No. 4.2e-134;			
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;	
QY	1	ATGAACCGGGAGTCCCTTTTAGGCACTGCTTCGCGTCGCAACTGGGCTCTCCCA	60		
Db	76	ATGAACCGGGAGTCCCTTTTAGGCACTGCTTCGCGTCGCAACTGGGCTCTCCCA	135		
QY	61	GCAGCACTCAGGAGAAAGATGTGCTGGGCAAAAAAGGAGATACAGTGAACCTGACC	120		
Db	136	GCAGCACTCAGGAGAAACAAAGTGTGCTGGGCAAAAAAGGAGATACAGTGAACCTGACC	195		
QY	121	TGTACAGTTCCCAAGAAAGCATACATTCCACTGGAAAACTCCACCAATTAAG	180		
Db	196	TGTACAGTTCCCAAGAAAGCATACATTCCACTGGAAAACTCCACCAATTAAG	255		
QY	181	ATTTCGGAAATCAGGGCTCCCTCTTAATAAGGTCATCCAAAGCTGAATGATCGGCT	240		
Db	256	ATTTCGGAAATCAGGGCTCCCTCTTAATAAGGTCATCCAAAGCTGAATGATCGGCT	315		
QY	241	GACTCAAGAAAGGCTTTGGGAGCCAAAGAACTTCCCTGATCATCAAGAACTTTAAG	300		
Db	316	GACTCAAGAAAGGCTTTGGGAGCCAAAGAACTTCCCTGATCATCAAGAACTTTAAG	375		
QY	301	ATAGAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG	360		
Db	376	ATAGAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG	435		
QY	361	CTAGTGTCGGAATTGACCTGCCAATCTTGACACCCACTGCTTCAAGGGCAGAGCTGACC	420		
Db	436	CTAGTGTCGGAATTGACCTGCCAATCTTGACACCCACTGCTTCAAGGGCAGAGCTGACC	495		
QY	421	CTGACCTTTGAGAGGCCCCCTGTGATGATGACCCCTCAGTGCATATGAGAGTCCAAAGGGT	480		
Db	496	CTGACCTTTGAGAGGCCCCCTGTGATGATGACCCCTCAGTGCATATGAGAGTCCAAAGGGT	555		
QY	481	AAAAACATACAGGGGGGGAAGACCTCTTCGCTGTCACTGAGAGCTTCAGGATATGGC	540		
Db	556	AAAAACATACAGGGGGGGAAGACCTCTTCGCTGTCACTGAGAGCTTCAGGATATGGC	615		
QY	541	ACCTGGAATGCACTGCTTGCAAGAACCAAGAAAGGTGGAGTTCAAAATGACATGGTG	600		
Db	616	ACCTGGAATGCACTGCTTGCAAGAACCAAGAAAGGTGGAGTTCAAAATGACATGGTG	675		

[illegible]

Query Match	43.2%;	Score 622.8;	DB 6;	Length 1273;
Best Local Similarity	99.7%;	Pred. No. 4,2e-134;		
Matches	624;	Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
QY	1	ATGAACCGGGGAGTCCCTTTTAAAGCACTGCTTGTGTCTGCAACTGGCGCTCCTCCCA	60	
Db	76	ATGAACCGGGGAGTCCCTTTTAAAGCACTGCTTGTGTGTCTGCAACTGGCGCTCCTCCCA	135	
QY	61	GCAGCCACTCAGGGGAAAGAAAGTGTGCTGGGCAAAAAGGGGATAcAGTGAACtGACC	120	
Db	136	GCAGCCACTCAGGGGAAAGAAAGTGTGTGCTGGGCAAAAAGGGGATACAGTGAACtGACC	195	
QY	121	TGTACAGCTTCCAGAAAGAAAGAGCATCAATTCCACTGGAATAAATCCTCAACAGATAAAG	180	
Db	196	TGTACAGCTTCCAGAAAGAAAGAGCATCAATTCCACTGGAATAAATCCTCAACAGATAAAG	255	
QY	181	ATTCTGGGAAATCAGGGGCTCCTTTTAACTTAAAGTCCATCCAAAGCTGAATGATCGCGCT	240	
Db	256	ATTCTGGGAAATCAGGGGCTCCTTTTAACTTAAAGTCCATCCAAAGCTGAATGATCGCGCT	315	
QY	241	GACTCAAGAAAGAAAGCTTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAAG	300	
Db	316	GACTCAAGAAAGAAAGCTTTTGGGACCAAGAAACCTTCCCTGATCATCAAGAACTTTAAAG	375	
QY	301	ATGAGAAGCTCAGATACTTACATCTGTGAAGTGAAGAACCAAGAGAGAGGTGCATTG	360	
Db	376	ATGAGAAGCTCAGATACTTACATCTGTGAAGTGAAGAACCAAGAGAGAGGTGCATTG	435	
QY	361	CTAGTGTGGGATTGACGCGCAACTGCAACCCACCTGCTTCAAGGGGCAAGCGCTGACC	420	
Db	436	CTAGTGTGGGATTGACGCGCAACTGCAACCCACCTGCTTCAAGGGGCAAGCGCTGACC	495	
QY	421	CTGACCTTGGAGAGCCGCCCTGTAGTATGAGCCCTCAGTGCATATGTAGAGTCCAAGGGT	480	
Db	496	CTGACCTTGGAGAGCCGCCCTGTAGTATGAGCCCTCAGTGCATATGTAGAGTCCAAGGGT	555	
QY	481	AAAAAACATACAGGGGGGGGAAAGACCTCTCCGCTGTCTCAGCTGGAGCTCCAGATATGTGGC	540	
Db	556	AAAAAACATACAGGGGGGGGAAAGACCTCTCCGCTGTCTCAGCTGGAGCTCCAGATATGTGGC	615	
QY	541	ACCTGGAATGACATGCTCTTGTCAAGAACAGAAAGAGTGAATTCAAAATATGACATCGTG	600	
Db	616	ACCTGGAATGACATGCTCTTGTCAAGAACAGAAAGAGTGAATTCAAAATATGACATCGTG	675	
QY	601	GTGCTAGCTTTCCGAAGAGGCTCCGG	626	
Db	676	GTGCTAGCTTTCCGAAGAGGCTCCAG	701	

RESULT 13  
 LOCUS AR067943 1304 bp DNA linear PAT 29-SEP-1999  
 DEFINITION Sequence 28 from patent US 5851828.  
 ACCESSION AR067943  
 VERSION AR067943.1 GI:5999165  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 AUTHORS 1 (bases 1 to 1304)  
 TITLE Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
 JOURNAL Targeted cytolysis of HIV-infected cells by chimeric CD4  
 FEATURES receptor-bearing cells  
 Patent: US 5851828-A 28 22-DEC-1998;  
 Location/Qualifiers  
 1..1304  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-134;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGCGCTCTCCCA 60  
 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGCGCTCTCCCA 170  
 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 120  
 171 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 230  
 121 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAATACTCCACAGATTAAG 180  
 231 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAATACTCCACAGATTAAG 290  
 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGTGAATGATCGGCT 240  
 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGTGAATGATCGGCT 350  
 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 351 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 410  
 301 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGTGAATG 360  
 411 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGTGAATG 470  
 361 CTAGTGTGGATTTGACTGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 420  
 471 CTAGTGTGGATTTGACTGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 530  
 421 CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGAATGTAGAGATCCAAAGGGT 480  
 531 CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGAATGTAGAGATCCAAAGGGT 590  
 481 AAAAATATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATATGAGC 540  
 591 AAAAATATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATATGAGC 650  
 541 ACCTGACATGCACTGTCTTTCAGAACCAAGAAAGTGGAGTTCAAAATAGATGTG 600  
 651 ACCTGACATGCACTGTCTTTCAGAACCAAGAAAGTGGAGTTCAAAATAGATGTG 710  
 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 14

AR166802  
 LOCUS AR166802 1304 bp DNA linear PAT 17-OCT-2001  
 DEFINITION Sequence 28 from patent US 6284240.  
 ACCESSION AR166802  
 VERSION AR166802.1 GI:16243143  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 REFERENCE Unclassified.  
 AUTHORS 1 (bases 1 to 1304)  
 TITLE Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
 JOURNAL Targeted cytolysis of HIV-infected cells by chimeric CD4  
 FEATURES receptor-bearing cells  
 Patent: US 6284240-A 28 04-SEP-2001;  
 Location/Qualifiers  
 1..1304  
 /organism="unknown"  
 /mol\_type="unassigned DNA"

ORIGIN  
 Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 4.2e-134;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGCGCTCTCCCA 60  
 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGTGGTGGCAACTGGCGCTCTCCCA 170  
 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 120  
 171 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAAGGGATACAGTGAATGACC 230  
 121 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAATACTCCACAGATTAAG 180  
 231 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAATACTCCACAGATTAAG 290  
 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGTGAATGATCGGCT 240  
 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGTGAATGATCGGCT 350  
 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
 351 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 410  
 301 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGTGAATG 360  
 411 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAAAGAGTGAATG 470  
 361 CTAGTGTGGATTTGACTGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 420  
 471 CTAGTGTGGATTTGACTGCAACTGTGACACCACTGCTTCAGGGGCAAGCTTACC 530  
 421 CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGAATGTAGAGATCCAAAGGGT 480  
 531 CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGAATGTAGAGATCCAAAGGGT 590  
 481 AAAAATATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATATGAGC 540  
 591 AAAAATATACAGGGGGGAAAGCCCTCTCGTGTCTCAGTGAAGCTCCAGATATGAGC 650  
 541 ACCTGACATGCACTGTCTTTCAGAACCAAGAAAGTGGAGTTCAAAATAGATGTG 600  
 651 ACCTGACATGCACTGTCTTTCAGAACCAAGAAAGTGGAGTTCAAAATAGATGTG 710  
 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 15  
 AR560088 1304 bp mRNA linear PAT 08-OCT-2004  
 LOCUS AR560088  
 DEFINITION Sequence 28 from patent US 6753162.

ACCESSION ARS60088  
VERSION ARS60088.1 GI:53970428  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1304)  
AUTHORS Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
receptor-bearing cells  
JOURNAL Patent: US 6753162-A 28 22-JUN-2004;  
FEATURES location/Qualifiers  
source 1..1304  
/organism="unknown"  
/mol\_type="mRNA"

## ORIGIN

Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 4.2e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAAACCGGGAGTCCCTTTAGCACTTCTGCTGCAACTGGCGTCTCCCA 60  
DB |||||  
QY 111 ATGAACCGGGAGTCCCTTTAGCACTTCTGCTGCAACTGGCGTCTCCCA 170  
DB |||||  
QY 61 GCAGCCACTAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACATGAACTGACC 120  
DB |||||  
QY 171 GCAGCCACTAGGAAAGAAAGTGTGCTGGCAAAAAGGGGATACATGAACTGACC 230  
DB |||||  
QY 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCATGGAATACTCAACCAATAAG 180  
DB |||||  
QY 231 TGTACAGCTTCCCAAGAAAGAGATCAATTCATGGAATACTCAACCAATAAG 290  
DB |||||  
QY 181 ATTCTGGGAATCAGGCTCTTCTTAATAAGTCCATCAAGTGAATGCGGCT 240  
DB |||||  
QY 291 ATTCTGGGAATCAGGCTCTTCTTAATAAGTCCATCAAGTGAATGCGGCT 350  
DB |||||  
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTAAG 300  
DB |||||  
QY 351 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTAAG 410  
DB |||||  
QY 301 ATGAAGACTCAGTACTTCACTGTGAGTGGAGCAAGAGGAGGTCATTTG 360  
DB |||||  
QY 411 ATGAAGACTCAGTACTTCACTGTGAGTGGAGCAAGAGGAGGTCATTTG 470  
DB |||||  
QY 361 CTAGTGTTCGGAATTGACTGCAACTTGACACCACTGCTTCAAGGCGAGGCTGACC 420  
DB |||||  
QY 471 CTAGTGTTCGGAATTGACTGCAACTTGACACCACTGCTTCAAGGCGAGGCTGACC 530  
DB |||||  
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCTCAGTCAATGTAGAGTCCAAAGGGT 480  
DB |||||  
QY 531 CTGACCTTGGAGAGCCCCCTGTAGTACCCTCAGTCAATGTAGAGTCCAAAGGGT 590  
DB |||||  
QY 481 AAAAACAATAAGGGGGGAGAACCCCTCCGTGTCAAGTGGAGCTCCAGGATAGTGGC 540  
DB |||||  
QY 591 AAAAACAATAAGGGGGGAGAACCCCTCCGTGTCTCAGCTGGAGCTCCAGGATAGTGGC 650  
DB |||||  
QY 541 ACCTGACATGCACTGTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
DB |||||  
QY 651 ACCTGACATGCACTGTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 710  
DB |||||  
QY 601 GTGCTAGCTTTCCAGAAAGGCTTCGG 626  
DB |||||  
QY 711 GTGCTAGCTTTCCAGAAAGGCTTCAG 736  
DB |||||

Search completed: September 20, 2005, 10:17:33  
Job time : 6463 secs



PT HIV neutralization, comprises two different binding domains, inducing-  
 PT binding domain and induced-binding domain functionally linked by linker.  
 PS Claim 36; Page 47-48; 55pp; English.  
 XX

CC sCD4-sCFv(17b) is a neutralizing bispecific fusion protein capable of  
 CC binding to two sites of its target protein. The protein comprises a first  
 CC binding domain capable of binding to an inducing site on the target  
 CC protein, a second binding domain capable of forming neutralising complex  
 CC with an induced epitope of the target protein and a linker connecting the  
 CC (binding domains. sCD4-sCFv(17b) comprises a soluble CD4 fragment  
 CC (containing domains D1 and D2) fused to a single chain Fv portion of  
 CC antibody 17b via a linker. sCD4-sCFv(17b), its variant, analogue or  
 CC derivative is used for inactivating gp120 protein of HIV, and for  
 CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fe-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure  
 XX

80 Sequence 1440 BP; 345 A; 354 C; 452 G; 289 T; 0 U; 0 Other;

Query Match 100.0%; Score 1440; DB 3; Length 1440;

Best Local Similarity 100.0%; Pred. No. 1.5e-311;

Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAAACCGGGGAGTCTTTTAAAGCATTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 1 ATGAAACCGGGGAGTCTTTTAAAGCATTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 120  
 DB 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACGTAGC 120  
 QY 121 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGGAATACTCCAAACGATTAAG 180  
 DB 121 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGGAATACTCCAAACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGTCAATCCAGCTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGTCAATCCAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCGCTGATCATCAAAATCTTAAG 300  
 DB 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCGCTGATCATCAAAATCTTAAG 300  
 QY 301 ATAGAAAGCTCAGATTAATCTGTAAGTGAAGAGCAAGAAAGAGAGGTGCAATTG 360  
 DB 301 ATAGAAAGCTCAGATTAATCTGTAAGTGAAGAGCAAGAAAGAGAGGTGCAATTG 360  
 QY 361 CTAGTGTTCGGATTGATGCAACTGCAACCACTGCTTCAAGGGGAGAGCTTAAC 420  
 DB 361 CTAGTGTTCGGATTGATGCAACTGCAACCACTGCTTCAAGGGGAGAGCTTAAC 420  
 QY 421 CTGACCTTTGGAGAGCCCCCTGTTAGTACCCCTCAATGCAATGAGATCCAAAGGGGT 480  
 DB 421 CTGACCTTTGGAGAGCCCCCTGTTAGTACCCCTCAATGCAATGAGATCCAAAGGGGT 480  
 QY 481 AAAAACAATACAGGGGGGAAACCTCTCCGTGTTCTGAGTGAAGCTCAAGATAGTGC 540  
 DB 481 AAAAACAATACAGGGGGGAAACCTCTCCGTGTTCTGAGTGAAGCTCAAGATAGTGC 540  
 QY 541 ACCTGGACATGCACTGTTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATGACATCTG 600

DB 541 ACCTGGACATGCACTGTTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATGACATCTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGCAGTATGGGGAGGCGGTTCCAGCGGA 660  
 DB 601 GTGCTAGCTTTCCAGAAAGGCTCCGAGAGTGCAGTATGGGGAGGCGGTTCCAGCGGA 660  
 QY 661 GGTGATCCGGTGGCGGAGAGGTTCGGCGGGGGGTGAAGCGGGGGTGGCGGCTCCGAGGC 720  
 DB 661 GGTGATCCGGTGGCGGAGAGGTTCGGCGGGGGGTGAAGCGGGGGTGGCGGCTCCGAGGC 720  
 QY 721 GGAAGTTCAAGGTGAGTGTCTCCAGTCTGGGGCTGAGAGTGAAGAGCTGGTCTCTG 780  
 DB 721 GGAAGTTCAAGGTGAGTGTCTCCAGTCTGGGGCTGAGAGTGAAGAGCTGGTCTCTG 780  
 QY 781 GTAAGGTCTCCGCAAGGCTCTGAGAGACCTTCATCAGATATATGTTACTGGGTG 840  
 DB 781 GTAAGGTCTCCGCAAGGCTCTGAGAGACCTTCATCAGATATATGTTACTGGGTG 840  
 QY 841 CGACAGGCTCTGAGCAAGGCTTGAATGATGGAGAGATCATCATCTTGAATGA 900  
 DB 841 CGACAGGCTCTGAGCAAGGCTTGAATGATGGAGAGATCATCATCTTGAATGA 900  
 QY 901 GCACACTAGCAACCGGACCTTCAGGGCAGAGTCAAGATTAACCGGAGCAAGTCCAGAC 960  
 DB 901 GCACACTAGCAACCGGACCTTCAGGGCAGAGTCAAGATTAACCGGAGCAAGTCCAGAC 960  
 QY 961 ACAATCTACCTGAGGTGCGGAATCTAAGATCTGACATACGAGCGGTATATTTCTGTGCG 1020  
 DB 961 ACAATCTACCTGAGGTGCGGAATCTAAGATCTGACATACGAGCGGTATATTTCTGTGCG 1020  
 QY 1021 GGAATCTAGAGGAGAGCGAGCGAGCAAGGAGATATGATTAATGAGTTCTGAAACAT 1080  
 DB 1021 GGAATCTAGAGGAGAGCGAGCGAGCAAGGAGATATGATTAATGAGTTCTGAAACAT 1080  
 QY 1081 TGGGGCCAGGGAACCTGTGTCAGGTCACTGAGGTGGGGTCCGAGAGTGGTGG 1140  
 DB 1081 TGGGGCCAGGGAACCTGTGTCAGGTCACTGAGGTGGGGTCCGAGAGTGGTGG 1140  
 QY 1141 AGCGGTGGCGGAGATCTGAACTGAGTTGAGCAGTCTCCAGCACCTGTCTGTGCT 1200  
 DB 1141 AGCGGTGGCGGAGATCTGAACTGAGTTGAGCAGTCTCCAGCACCTGTCTGTGCT 1200  
 QY 1201 CCAGGGGAAAGAGCACCTCTCTGCAAGGCGCAGTGAAGTGTAGTGAAGCTTAAGC 1260  
 DB 1201 CCAGGGGAAAGAGCACCTCTCTGCAAGGCGCAGTGAAGTGTAGTGAAGCTTAAGC 1260  
 QY 1261 TGGTACACAGAAACCTGGCCAGAGCTCCAGGCTCCTCATATATGATGATCCACAGG 1320  
 DB 1261 TGGTACACAGAAACCTGGCCAGAGCTCCAGGCTCCTCATATATGATGATCCACAGG 1320  
 QY 1321 GCCACCGGTGTCACAGCAGGTTCAAGTGGAGTGGGGTGGGGCAGAAATTCATCTAC 1380  
 DB 1321 GCCACCGGTGTCACAGCAGGTTCAAGTGGAGTGGGGTGGGGCAGAAATTCATCTAC 1380  
 QY 1381 ATCAGCAGCTGCAAGTCTGAAGATTTTGCAGTTTATTACTGTACAGCATTAATCTGG 1440  
 DB 1381 ATCAGCAGCTGCAAGTCTGAAGATTTTGCAGTTTATTACTGTACAGCATTAATCTGG 1440

# RESULT 2

ADN07735 standard; DNA; 8911 BP.

ADN07735;

15-JUL-2004 (first entry)

Expression vector DNA.

Immunogenic complex; gp120; CD4; HIV; HIV infection; anti-HIV; circular;

cyclic; ds.

XX	SYNTHETIC.	
XX	US2004076636-A1.	
XX	22-APR-2004.	
XX	02-JUL-2003; 2003US-00612192.	
XX	07-MAY-1993; 93US-00060926.	
XX	06-MAY-1994; 94WO-US005020.	
XX	20-DEC-1995; 95US-00464680.	
XX	11-MAY-1998; 98US-00075544.	
XX	17-JUN-2000; 2000US-00479675.	
XX	17-JUN-2001; 2001US-00905962.	
XX	(PALR/) PAL R.	
XX	(MARK/) MARKHAM P.	
XX	(KEEN/) KEEN T.	
XX	(WHIT/) WHITNEY S.	
XX	(KALY/) KALYANARAMAN V S.	
XX	Pal R, Markham P, Keen T, Whitney S, Kalyanaraman VS;	
XX	WPI; 2004-387924/36.	
XX		
PT	Immunogenic complex, useful for treating HIV infections, comprises gp120	
PT	covalently bonded to CD4 equivalent molecule, fragment of CD4 or its	
PT	equivalent.	
XX		
XX	Disclosure; SEQ ID NO 3; 26pp; English.	
XX		
CC	The invention relates to an immunogenic complex comprising gp120	
CC	covalently bonded to a CD4 equivalent molecule, fragment of CD4 or its	
CC	equivalent. The invention also relates to a composition comprising the	
CC	complex and a carrier, an antibody reactive with the complex, an	
CC	immortalised cell line that produces the complex, a method of detecting	
CC	the HIV antigen in a test fluid, involving contacting the test fluid with	
CC	an antibody raised against the immunogenic complex and detecting the	
CC	presence of immune complexes formed between the antigen in the test fluid	
CC	and the antibody, and a vaccine comprising an immunogenically effective	
CC	amount of the immunogenic complex. The immunogenic complex is useful for	
CC	raising neutralising antibodies against HIV, which involves administering	
CC	the complex to a subject in a carrier, and for treating HIV infections.	
CC	This sequence represents expression vector DNA used in the scope of the	
CC	invention.	
CC		
CC	Sequence 8911 BP; 2188 A; 2221 C; 2227 G; 2275 T; 0 U; 0 Other;	
SO		
Query Match	43.4%; Score 625; DB 12; Length 8911;	
Best Local Similarity	100.0%; Pred. No. 1.9e-129;	
Matches 625; Conservative	0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGAACCGGGAGTCCCTTTTAGGACCTGCTTCGTGTCGCACTGCGCTCTCCCA	60
DB	3815 ATGAACCGGGAGTCCCTTTTAGGACCTGCTTCGTGTCGCACTGCGCTCTCCCA	3874
QY	61 GCAGCACTCAGGAGAAAGTGTGTGCGCAAAAAAGGGATACAGTGAC	120
DB	3875 GCAGCACTCAGGAGAAAGTGTGTGCGCAAAAAAGGGATACAGTGAC	3933
QY	121 TGTACAGCTTCCAGAGAGACATACATTCACCTGAAAACTCCAACGATTAAG	180
DB	3935 TGTACAGCTTCCAGAGAGACATACATTCACCTGAAAACTCCAACGATTAAG	3994
QY	181 ATTCTGGGAATCAGGCGCTCTTTAATTAAGTCCATCCAACTGAATATGCGCT	240
DB	3995 ATTCTGGGAATCAGGCGCTCTTTAATTAAGTCCATCCAACTGAATATGCGCT	4053
QY	241 GACTCAAGAAAGCTTTTGGGACCAAGAACTTCCCTGATATCAAGATCTTAAG	300
DB	4055 GACTCAAGAAAGCTTTTGGGACCAAGAACTTCCCTGATATCAAGATCTTAAG	4111
QY	301 ATGAGAAGCTCAGATATTACATCTGTGAATGAGAGACAGAGAGAGGTGCAATTG	360

Db	4115	ATTAAGACTCAGTACTTTCATCTGTGAAGTGAGAGCAAGAAAGAGAGTGCAATTG	4174
Oy	361	CTAGTGTTCGAGATTGACTGCCAACTGTGACACCACTGCTTCAGGGGCAAGCCTGACC	420
Db	4175	CTAGTGTTCGAGATTGACTGCCAACTGTGACACCACTGCTTCAGGGGCAAGCCTGACC	4234
Oy	421	CTGACCTTTGAGAGCCCCCTGGTAGTAGGCCCTCAGTGCATATGTATGAGTCCAAAGGGGT	480
Db	4235	CTGACCTTTGAGAGCCCCCTGGTAGTAGGCCCTCAGTGCATATGTATGAGTCCAAAGGGGT	4294
Oy	481	AAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATAGTGC	540
Db	4295	AAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATAGTGC	4354
Oy	541	ACCTGACATGACATGCTGCTTTGCAGAACCAAGAGGTGAGTTCAAAATAGACATCTGTG	600
Db	4355	ACCTGACATGACATGCTGCTTTGCAGAACCAAGAGAGGTGAGTTCAAAATAGACATCTGTG	4414
Oy	601	GTCCTAGCTTTCCAGAAAGGCTCCG	625
Db	4415	GTCCTAGCTTTCCAGAAAGGCTCCG	4439

RESULT 3

AAF82582

ID AAF82582 standard; cDNA; 1377 BP.

XX AAF82582;

XX AC

XX

XX 18-JUN-2001 (first entry)

XX DE

Human CD4 gene T4.

XX

Human; CD4; T4; CD4 fusion protein; oligomerisation;

KW receptor-ligand interaction inhibition; surface plasmon resonance; SPR;

KW T cell receptor binding; MHC binding; carcinoma; autoimmune disease;

KW multiple sclerosis; human immunodeficiency virus; HIV; diabetes;

KW rheumatoid arthritis; immune disorder; ss.

XX

OS Homo sapiens.

XX

XX

Key Location/Qualifiers

FT CDS 1..1377

FT /\*tag= a

FT /product= "human CD4"

FT sig\_peptide 1..75

FT /\*tag= b

FT mat\_peptide 76..1374

FT /\*tag= c

XX

XX MO20012084-A2.

XX

XX 29-MAR-2001.

XX

XX 18-SEP-2000; 2000MO-GB003579.

XX

XX 21-SEP-1999; 99GB-00022352.

XX

XX (AVID-) AVIDEX LTD.

XX

XX Jakobson BK;

XX

XX WPI: 2001-273470/28.

XX

XX P-RSDB: AAB01502.

XX

XX Sequential screening of candidate compounds library for those which

XX inhibit binding of low affinity receptor-ligand interaction having fast

XX binding kinetics, using interfacial optical assay.

XX

XX Disclosure; Fig 13; 91pp; English.

XX

XX The present sequence encodes human CD4. Human CD4 extracellular domains 1

CC and 2 were used in the construction of CD4 oligomerisation fusion  
 CC proteins. The fusion proteins contain an oligomerisation domain that  
 CC enables the proteins to bind to one another to form oligomers. The  
 CC oligomers may be used in an invention relating to a method for screening  
 CC for compounds with the ability to inhibit a low affinity receptor-ligand  
 CC interaction. The method uses an interfacial optical assay, such as  
 CC surface plasmon resonance (SPR). The method is useful for screening  
 CC candidate compounds for the ability to inhibit interaction between  
 CC MHC/peptide complex and T cell receptor, and MHC/peptide complex and CD8  
 CC or CD4 co-receptor. The compounds identified by the above methods which  
 CC interfere with T cell receptor binding to a particular HLA type molecule  
 CC are useful as immune inhibitors for treating carcinoma, autoimmune  
 CC diseases such as multiple sclerosis, human immunodeficiency virus (HIV)  
 CC infection, rheumatoid arthritis, Hashimoto's disease, insulin dependent  
 CC diabetes, Good pasture's syndrome, uveitis, psoriasis and graft rejection  
 CC  
 XX  
 SQ Sequence 1377 BP; 351 A; 355 C; 393 G; 278 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 4; Length 1377;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGAGCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 1 ATGACCGGGAGAGCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAGGGATACATGTAATGACACC 120  
 DB 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAGGGATACATGTAATGACACC 120  
 QY 121 TGTACAGCTTCCAGAAAGAGATATTCATTCCTGAGAAAACCTCCAACTGATTAAG 180  
 DB 121 TGTACAGCTTCCAGAAAGAGATATTCATTCCTGAGAAAACCTCCAACTGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCCAGCTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCCAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 241 GACTCAAGAAAGAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 QY 301 ATGAGAACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 360  
 DB 301 ATGAGAACTCAATATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCATTTG 360  
 QY 361 CTAGTGTTCGGAATGACTGCAACTCTGACCACTGCTTCAAGGGGCAAGGCTGACC 420  
 DB 361 CTAGTGTTCGGAATGACTGCAACTCTGACCACTGCTTCAAGGGGCAAGGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATATGTAGAGTCCAAAGGAT 480  
 DB 421 CTGACCTTGGAGAGCCCTCTGTGTAGTAGCCCTCAGTGCATATGTAGAGTCCAAAGGAT 480  
 QY 481 AAAAATCTAAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
 DB 481 AAAAATCTAAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
 QY 541 ACCGTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAAGTTCAAAATAGACATGTTG 600  
 DB 541 ACCGTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAAGTTCAAAATAGACATGTTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

RESULT 4  
 ID AA038761 standard; DNA; 1415 BP.  
 XX  
 AC AA038761;  
 XX

DT 25-MAR-2003 (revised)  
 DT 22-JUL-1993 (first entry)  
 XX  
 XX sCD4-L2 lysosomal targeting fusion gene.  
 DE  
 XX Soluble CD4; HIV; Human immunodeficiency Virus; envelope; glycoprotein;  
 KM polymerase chain reaction; lysosomal membrane protein; Lamp-2; 86.  
 XX  
 XX Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT misc\_feature 1..6  
 FT /tag= a  
 FT /note= "EcoRI restriction site"  
 FT misc\_feature 1275..1280  
 FT /tag= b  
 FT /note= "XbaI restriction site"  
 FT misc\_feature 1410..1415  
 FT /tag= c  
 FT /note= "SalI restriction site"  
 FT  
 PN MO9306216-A1.  
 XX  
 PD 01-APR-1993.  
 XX  
 PF 22-SEP-1992; 92WO-US008090.  
 XX  
 PR 26-SEP-1991; 91US-00766963.  
 XX  
 PA (OKLA-) OKLAHOMA MED RES FOUND.  
 XX  
 PI Tang JUN, Lin XL;  
 XX  
 DR WPI; 1993-117537/14.  
 XX  
 PT New fusion protein used in gene therapy for treating AIDS - comprises  
 PT protein which binds to retroviral envelope protein which targets fusion  
 PT protein to lysosome.  
 PS  
 XX Claim 12; Page 14-15; 47pp; English.  
 CC  
 CC Primers P-1 and P-2 (AA038748 and AA038749) were used to amplify an EcoRI  
 CC -XbaI fragment of plasmid p14B containing the sCD4 sequence. The PCR  
 CC product was ligated to an XbaI-SalI fragment containing the sequence  
 CC encoding lysosomal membrane protein (LAMP-2) (PCR amplified from a human  
 CC liver lambda gp10 library using primers AA038756 and AA038757). The  
 CC resulting fusion sequence sCD4-L2 can be used to interfere with the  
 CC normal function of HIV and to direct the virus to lysosomes. (Updated on  
 CC 25-MAR-2003 to correct PN field.)  
 XX  
 SQ Sequence 1415 BP; 360 A; 365 C; 393 G; 297 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1415;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGAGCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 81 ATGACCGGGAGAGCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCTCCCA 140  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAGGGATACATGTAATGACACC 120  
 DB 141 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAGGGATACATGTAATGACACC 200  
 QY 121 TGTACAGCTTCCAGAAAGAGATATTCATTCCTGAGAAAACCTCCAACTGATTAAG 180  
 DB 201 TGTACAGCTTCCAGAAAGAGATATTCATTCCTGAGAAAACCTCCAACTGATTAAG 260  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCCAGCTGAATGATCGGCT 240  
 DB 261 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCCAGCTGAATGATCGGCT 320  
 QY 241 GACTCAAGAAAGAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300

DB 321 GACTCAAGAAAGCCTTTGGAGCAAGAACTCCCTGATCATCAAGAACTTAAG 380  
 QY 301 ATAGAACTCAGATTAATCATCTGTGAAGTGGAGCAGAGAGAGAGTCAATTG 360  
 DB 381 ATAGAACTCAGATTAATCATCTGTGAAGTGGAGCAGAGAGAGAGTCAATTG 440  
 QY 361 CTAGTGTGGATTAATCATCTGTGAAGTGGAGCAGAGAGAGTCAATTG 420  
 DB 441 CTAGTGTGGATTAATCATCTGTGAAGTGGAGCAGAGAGAGTCAATTG 500  
 QY 421 CTGACCTTGGAGAGCCCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 480  
 DB 501 CTGACCTTGGAGAGCCCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 560  
 QY 481 AAAAATACAGAGGGGGGAGAGACCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 540  
 DB 561 AAAAATACAGAGGGGGGAGAGACCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 620  
 QY 541 ACCTGACATGACATCTGTCTTGCAGAACCAAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 DB 621 ACCTGACATGACATCTGTCTTGCAGAACCAAGAGAGTGAAGTCAAAATAGACATGCTG 680  
 QY 601 GTGCTAGCTTTCCAGAAAGCCTCTCG 626  
 DB 681 GTGCTAGCTTTCCAGAAAGCCTCTCG 706  
 RESULT 5  
 ADA44806  
 ID ADA44806 standard; DNA; 1419 BP.  
 AC ADA44806;  
 XX  
 DT 04-DEC-2003 (first entry)  
 DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon1on15 DNA, SEQ ID NO:1.  
 XX  
 KM HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;  
 KM endoplasmic reticulum; ER retention; envelope protein gp160;  
 KM T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon1on15;  
 KM gene therapy; human; gene; ds.  
 XX  
 OS Chimeric.  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1419  
 FT /\*tag= a  
 FT /partial  
 FT /product= "CD4epsilon1on15"  
 FT /note= "No stop codon given"  
 FT  
 PN WO2003076468-A1.  
 PD 18-SEP-2003.  
 XX  
 PF 14-MAR-2003; 2003WO-ES000120.  
 XX  
 PR 14-MAR-2002; 2002ES-0000616.  
 XX  
 PA (NNSU) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
 PI Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
 XX Gomez Buendia M;  
 DR WPI; 2003-779059/73.  
 DR P-PSDB; ADA44807.  
 XX  
 PT Composition for treating or preventing human immune deficiency virus,  
 PT comprises CD4 chimeric protein having a protective effect in trans, or  
 PT related nucleic acid.  
 XX

PS Claim 5; Page 31-33; 43pp; Spanish.  
 XX  
 CC The invention relates to a composition for the treatment or prevention of  
 CC human immunodeficiency virus-1 (HIV-1) infection. The composition  
 CC comprises CD4+ cells that have been transduced with a vector that encodes  
 CC a chimeric CD4 molecule which is capable of being retained in the  
 CC endoplasmic reticulum (ER). The invention also encompasses the use of a  
 CC soluble protein factor produced by CD4+ cells that have been transduced  
 CC with a vector encoding a chimeric CD4 protein, and the use of an  
 CC expression system encoding a chimeric CD4 protein. The ER-localised  
 CC chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
 CC resulting in HIV-1 retention in the ER and thereby preventing viral  
 CC replication. In a specific embodiment, the chimeric CD4 molecule  
 CC comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
 CC cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
 CC CD4epsilon1on15 (ADA44807). A known chimeric CD4 of similar structure but  
 CC containing only 10 amino acids from CD3epsilon can also be used.  
 CC Compositions of the invention have an in trans effect on the replication  
 CC of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
 CC present sequence represents a DNA encoding the chimeric CD4 molecule  
 CC CD4epsilon1on15, which is specifically claimed for use in compositions of  
 CC the invention.  
 XX  
 SQ Sequence 1419 BP; 362 A; 367 C; 405 G; 285 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 10; Length 1419;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCACTGCGCTCTCCCA 60  
 DB 1 ATGAAACCGGGAGATCCCTTTTAAAGCACTTGTGCTGCACTGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAGGGATACAGTGAACAGACC 120  
 DB 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAGGGATACAGTGAACAGACC 120  
 QY 121 TGTACAGCTTCCAGAAAGAGCATATCAATTCATGGAATACTCCAGATTAAG 180  
 DB 121 TGTACAGCTTCCAGAAAGAGCATATTCATGGAATACTCCAGATTAAG 180  
 QY 181 ATTCTGGGAATCAGAGGCTCTTCTTAATCAAGTTCATCAAGCTGAATGCGGCT 240  
 DB 181 ATTCTGGGAATCAGAGGCTCTTCTTAATCAAGTTCATCAAGCTGAATGCGGCT 240  
 QY 241 GACTCAAGAAAGAGCCTTTGGAGCAAGAGAACTTCCCTGATCAAGAAATCTTAAG 300  
 DB 241 GACTCAAGAAAGAGCCTTTGGAGCAAGAGAACTTCCCTGATCAAGAAATCTTAAG 300  
 QY 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAAGAGCAAGAGAGAGTGAATTG 360  
 DB 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAAGAGCAAGAGAGAGTGAATTG 360  
 QY 361 CTAGTGTGGATTAATCATCTGTGAAGTGGAGCAGAGAGAGTCAATTG 420  
 DB 361 CTAGTGTGGATTAATCATCTGTGAAGTGGAGCAGAGAGAGTCAATTG 420  
 QY 421 CTGACCTTGGAGAGCCCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 480  
 DB 421 CTGACCTTGGAGAGCCCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 480  
 QY 481 AAAAATACAGAGGGGGGAGAGACCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 540  
 DB 481 AAAAATACAGAGGGGGGAGAGACCTCTGAGTGAAGCCCTCAATGTAGAGTCCAGAGGCT 540  
 QY 541 ACCTGACATGACATCTGTCTTGCAGAACCAAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 DB 541 ACCTGACATGACATCTGTCTTGCAGAACCAAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGCCTCTCG 626  
 DB 601 GTGCTAGCTTTCCAGAAAGCCTCTCG 626

ID	AAQ38760	standard; DNA; 1421 BP.
AC	AAQ38760;	
DT	25-MAR-2003	(revised)
DT	22-JUL-1993	(first entry)
DB	scd4-l1 lysosomal targeting fusion gene.	
KW	Soluble CD4; HIV; Human immunodeficiency Virus; envelope; glycoprotein;	
KW	polymerase chain reaction; lysosomal membrane protein; lamp-1; ss.	
OS	Synthetic.	
FT	Key	Location/Qualifiers
FT	misc_feature	1..6
FT	misc_feature	/*tag= a
FT	misc_feature	/note= "EcoRI restriction site"
FT	misc_feature	1275..1280
FT	misc_feature	/*tag= b
FT	misc_feature	/note= "XbaI restriction site"
FT	misc_feature	1416..1421
FT	misc_feature	/*tag= c
FT	misc_feature	/note= "Sall restriction site"
PN	WO9306216-A1.	
PD	01-APR-1993.	
PF	22-SEP-1992;	92WO-US008090.
PR	26-SEP-1991;	91US-00766963.
PA	(OKLA-) OKLAHOMA MED RES FOUND.	
PI	Tang JN, Lin XL;	
PI	WPI; 1993-117537/14.	
PT	New fusion protein used in gene therapy for treating AIDS - comprises protein which binds to retroviral envelope protein which targets fusion protein to lysosome.	
PS	Claim 12; Page 13-14; 47pp; English.	
CC	Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI	
CC	-XbaI fragment of plasmid pT4B containing the scd4 sequence. The PCR	
CC	product was ligated to an XbaI-SalI fragment containing the sequence	
CC	encoding lysosomal membrane protein (lamp-1) (PCR amplified from a human	
CC	liver lambda gpt10 library using primers AAQ38754 and AAQ38755). The	
CC	resulting fusion sequence scd4-l1 can be used to interfere with the	
CC	normal function of HIV and to direct the virus to lysosomes. (Updated on	
CC	25-MAR-2003 to correct PN field.)	
CC	Sequence 1421 BP; 355 A; 382 C; 401 G; 283 T; 0 U; 0 Other;	

Query Match	Similarity	Score	DB 2:	Length
Best Local	99.8%	Pred. No. 1.8e-129;		
Matches	625;	Conservative	0;	Mismatches 1; Indels 0; Gaps 0;
QY	1	ATGAACCGGGAGATCCCTTTTAGGCACTTGCTCTGTGCTGCACCTGGCGTCTCTCCCA	60	
Db	81	ATGAACCGGGAGATCCCTTTTAGGCACTTGCTCTGTGCTGCACCTGGCGTCTCTCCCA	140	
QY	61	GCAGGCATCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC	120	
Db	141	GCAGGCATCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC	200	
QY	121	TGTACAGCTTCCCAAGAAAGCATCAATTCCACTGGAAAACTCCACCAAGTAAAG	180	

Db	201	TGTACGCTTCCAGAAAGAGCATACATTCACCTG3GAAAAAACTCCAAACGATTAAG	260
Oy	181	ATTCTGGGAAATCAGGGCTCTTCTTTAACTTAAGGTCATCAAGCTGAATGATCGCGCT	240
Db	261	ATTCTGGGAAATCAGGGCTCTTCTTTAACTTAAGGTCATCAAGCTGAATGATCGCGCT	320
Oy	241	GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGAACTTTAAG	300
Db	321	GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGAACTTTAAG	380
Oy	301	ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCGAATTG	360
Db	381	ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCGAATTG	440
Oy	361	CTAGTGTGGGATTTGATCTGCAACTCTGACACCCACCTGGCTTACGGGGACAGGCTGACC	420
Db	441	CTAGTGTGGGATTTGATCTGCAACTCTGACACCCACCTGGCTTACGGGGACAGGCTGACC	500
Oy	421	CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTTCAGTCAATGTAGAGATCCAAAGGGAT	480
Db	501	CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTTCAGTCAATGTAGAGATCCAAAGGGAT	560
Oy	481	AAAAAATATCAAGGGGGGGAAGACCTTCGCCGTCTCAGCTGGAGCTCCAGAGTAAGTGC	540
Db	561	AAAAAATATCAAGGGGGGGAAGACCTTCGCCGTCTCAGCTGGAGCTCCAGAGTAAGTGC	620
Oy	541	ACCTGGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATCGTG	600
Db	621	ACCTGGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATCGTG	680
Oy	601	GTCCTAGCTTTTCAAGAGGCTTCGGG	626
Db	681	GTCCTAGCTTTTCAAGAGGCTTCGGG	706

RESULT	7
AAQ38759	
ID	AAQ38759 standard; DNA; 1448 BP.
XX	
AC	AAQ38759;
XX	
D7	25-MAR-2003 (revised)
DT	22-JUL-1993 (first entry)
XX	
DE	sCD4-HAP lysosomal targeting fusion gene.
XX	
KW	Soluble CD4; HIV; Human immunodeficiency Virus; envelope; ss;
KW	glycoprotein; polymerase chain reaction; human acid phosphatase.
XX	
OS	Synthetic.
XX	
FH	Key
FT	misc_feature
FT	Location/Qualifiers
FT	1..6
FT	/tag= a
FT	/note= "ScorI restriction site"
FT	144..148
FT	/tag= b
FT	/note= "Sall restriction site"
XX	
PN	WO9306216-A1.
PD	
PD	01-APR-1993.
XX	
PP	22-SEP-1992; 92MO-US008090.
XX	
PR	26-SEP-1991; 91US-00766963.
PA	(OKLA-) OKLAHOMA MED RES FOUND.
PI	Tang JN, Lin XL;
DR	WPI, 1993-117537/14.

PT New fusion protein used in gene therapy for treating AIDS - comprises  
 PT protein which binds to retroviral envelope protein which targets fusion  
 PT protein to lysosome.

PS Claim 12; Page 12-13; 47pp; English.

CC Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI  
 CC -XbaI fragment of plasmid pT48 containing the scd4 sequence. The PCR  
 CC product was ligated to an XbaI-SalI fragment containing the sequence  
 CC encoding human acid phosphatase (HAP) lysosomal targeting protein (PCR  
 CC amplified from a human liver lambda gp10 library using primers AAQ38752  
 CC and AAQ38753). The resulting fusion sequence scd4-HAP can be used to  
 CC interfere with the normal function of HIV and to direct the virus to  
 CC lysosomes. (Updated on 25-MAR-2003 to correct PN field.)

XX Sequence 1448 BP; 357 A; 395 C; 404 G; 292 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1448;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 81 ATGAACCGGGAGTCCCTTTAGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 140  
 QY 61 GCAGCCACTCAGGAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120  
 DB 141 GCAGCCACTCAGGAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 200  
 QY 121 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGAAAACTCCAAACAGATAAG 180  
 DB 201 TGTACAGCTTCCAGAAAGAGCATTAATTCACCTGAAAACTCCAAACAGATAAG 260  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTTAAGTTCATTCAGTGAATGATCGGCT 240  
 DB 261 ATTCTGGGAATCAGGGCTCTCTTAACTTAAGTTCATTCAGTGAATGATCGGCT 320  
 QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTCCCGCTGATCATCAAGAACTTAAG 300  
 DB 321 GACTCAAGAAAGCTTTGGGACCAAGAACTCCCGCTGATCATCAAGAACTTAAG 380  
 QY 301 ATGAAGACTCAGTACTTATCATCTGTGAAGTGAAGCAAGAGAGAGTCAATTG 360  
 DB 381 ATGAAGACTCAGTACTTATCATCTGTGAAGTGAAGCAAGAGAGAGTCAATTG 440  
 QY 361 CTAGTGTTCGATGATCTGCAACTCTGACACCACTGCTTCAGGGGCAAGCTTACC 420  
 DB 441 CTAGTGTTCGATGATCTGCAACTCTGACACCACTGCTTCAGGGGCAAGCTTACC 500  
 QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCAATGTAGAGTCAAGGGGT 480  
 DB 501 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCAATGTAGAGTCAAGGGGT 560  
 QY 481 AAAAATACAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGAGC 540  
 DB 561 AAAAATACAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGCACTGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAATAATGACATGTG 600  
 DB 621 ACCTGCACTGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAATAATGACATGTG 680  
 QY 601 GTGCTAGCTTCCGAGAGGCTCCGG 626  
 DB 681 GTGCTAGCTTCCGAGAGGCTCCAG 706

RESULT 8  
 ID AAAS0662  
 AC AAAS0662; standard; DNA; 1714 BP.

XX 09-JAN-2001 (first entry)

XX DNA encoding CD4-IgM fusion protein CH4Mmu.  
 DE CD4; IgM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
 KW therapy; diagnosis; de.  
 KM  
 XX Homo sapiens.

OS	Key	Location/Qualifiers
XX	exon	1..1308
XX	exon	/*tag= b
XX	CDS	/number= 1
XX		111..1714
XX		/*tag= a
XX		/partial
XX	intron	/note= "contains an intron"
XX		1309..1664
XX		/*tag= c
XX	exon	/number= 1
XX		1665..1714
XX		/*tag= d
XX		/partial
XX		/number= 2

US6117656-A.  
 12-SEP-2000.  
 07-JUN-1995; 95US-00479353.  
 22-JAN-1988; 88US-00147351.  
 23-JAN-1989; 89US-00295956.  
 09-JUN-1992; 92US-0086781.  
 12-APR-1993; 93US-00057952.  
 04-FEB-1994; 94US-00191708.

PA (GENO ) GEN HOSPITAL CORP.

PI Seed B;

XX WPI; 2000-586558/55.

DR P-PSDB; AAB19509.

PT CD4-immunoglobulin fusion proteins, useful for targeting gp120 of HIV or  
 PT SIV.

XX Example 1; Col 41-50; 39pp; English.

XX The present sequence is that of DNA encoding fusion protein CD4Mmu (see  
 CC AAB19509) comprising the extracellular portion of CD4, which binds to HIV  
 CC gp120, linked at its C-terminus to a human IgM heavy chain polypeptide.  
 CC To obtain the construct, DNA encoding CD4 was linked to IgM DNA at the  
 CC Met2 site upstream of the CH1 region. A plasmid containing this genetic  
 CC construct is deposited in Escherichia coli MC1061/P3 as ATCC 67609.  
 CC Fusion protein CD4Mmu and a nucleic acid encoding it are claimed. Also  
 CC claimed are a vector comprising the nucleic acid, and a method of  
 CC producing the fusion protein in secreted form using a transformed host  
 CC cell. The fusion protein may further comprise a therapeutic agent,  
 CC radiolabel or NMR imaging agent. The fusion protein can be administered  
 CC to an animal (including humans) for treatment of HIV or SIV infection.  
 CC and can also be used in assays for HIV or SIV, imaging and tissue stains.  
 CC IgM fusion proteins such as CD4Mmu provide complement-mediated immunity  
 XX

SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTAGCACTTGCTTGTGCTGCAACTGGCGCTCTCCCA 170

```
QY 61 GCAGCCACTCAGGAGAAAGATGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 120
DB 171 GCAGCCACTCAGGAGAAAGATGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 230
QY 121 TGTACAGCTTCCAGAAAGAGATACATTCCTCACTGGAAGAAATCCCAACCAATTAAG 180
DB 231 TGTACAGCTTCCAGAAAGAGATACATTCCTCACTGGAAGAAATCCCAACCAATTAAG 230
QY 181 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGATCCATCAAGCTGAATGATCGGCT 240
DB 291 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGATCCATCAAGCTGAATGATCGGCT 350
QY 241 GACTCAAGAAAGAGCTTGGGAGCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300
DB 351 GACTCAAGAAAGAGCTTGGGAGCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 410
QY 301 ATGAGAACTCAATTAATTAATTTGTAAGTGAAGAGCAAGAAAGAGAGTGCATTTG 360
DB 411 ATGAGAACTCAATTAATTAATTTGTAAGTGAAGAGCAAGAAAGAGAGTGCATTTG 470
QY 361 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
DB 471 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 530
QY 421 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
DB 531 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
QY 481 AAAAACTAAGAGGGGGGAAAGCCCTCTCGTGTCTAAGTGAAGTCCAGATAGTGGC 540
DB 591 AAAAACTAAGAGGGGGGAAAGCCCTCTCGTGTCTAAGTGAAGTCCAGATAGTGGC 650
QY 541 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 600
DB 651 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 710
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 9
AAZ44063
ID AAZ44063 standard; DNA; 1714 BP.
XX
AC AAZ44063;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human fusion protein CD4Mg encoding DNA.
XX
KW Fusion protein; human; CD4; IgM; immunoglobulin; gp120;
XX anti-human immunodeficiency virus; CD4Mg; ds.
XX
OS Homo sapiens.
XX Synthetic.
XX
PN US6004781-A.
XX
PD 21-DEC-1999.
XX
PF 04-FEB-1994; 94US-00191708.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JUN-1989; 89US-00299596.
PR 03-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
AC WPI; 2000-085792/07.
XX
DR
```

```
DR P-PSDB; AAY51080.
XX
PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX
PS Example 1; Col 41-50; 39pp; English.
XX
CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence encodes
CC the fusion protein CD4Mg which is constructed from CD4 linked to human
CC IGM upstream of the CH1 region
XX
SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;
Query Match 43.4%; Score 624.4; DB 3; Length 1714;
Best Local Similarity 99.8%; Pred. No. 1.8e-129;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAAACCGGGAGATCCCTTTTAAAGCACTGCTTGTGCTGCTGCAACTGGCGCTCCCA 60
DB 111 ATGAAACCGGGAGATCCCTTTTAAAGCACTGCTTGTGCTGCTGCAACTGGCGCTCCCA 170
QY 61 GCAGCCACTCAGGAGAAAGATGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 120
DB 171 GCAGCCACTCAGGAGAAAGATGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 230
QY 121 TGTACAGCTTCCAGAAAGAGATACATTCCTCACTGGAAGAAATCCCAACCAATTAAG 180
DB 231 TGTACAGCTTCCAGAAAGAGATACATTCCTCACTGGAAGAAATCCCAACCAATTAAG 290
QY 181 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGATCCATCAAGCTGAATGATCGGCT 240
DB 291 ATTCTGGGAATCAGGGCTCTCTTCTTAATAAGATCCATCAAGCTGAATGATCGGCT 350
QY 241 GACTCAAGAAAGAGCTTGGGAGCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300
DB 351 GACTCAAGAAAGAGCTTGGGAGCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 410
QY 301 ATGAGAACTCAATTAATTAATTTGTAAGTGAAGAGCAAGAAAGAGAGTGCATTTG 360
DB 411 ATGAGAACTCAATTAATTAATTTGTAAGTGAAGAGCAAGAAAGAGAGTGCATTTG 470
QY 361 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 420
DB 471 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 530
QY 421 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 480
DB 531 CTAGTGTTCGGAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 590
QY 481 AAAAACTAAGAGGGGGGAAAGCCCTCTCGTGTCTAAGTGAAGTCCAGATAGTGGC 540
DB 591 AAAAACTAAGAGGGGGGAAAGCCCTCTCGTGTCTAAGTGAAGTCCAGATAGTGGC 650
QY 541 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 600
DB 651 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 710
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

RESULT 10
AAZ48203
ID AAZ48203 standard; DNA; 1714 BP.
XX
AC AAZ48203;
XX
DR
```

DT 14-MAR-2000 (first entry)  
 XX DNA sequence encoding CD4-Ig fusion protein CD4mmu.  
 DE  
 XX HIV, extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 KM secreted protein; SIV infection; medication; ds.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 PN CA1340741-C.  
 XX  
 PD 14-SEP-1999.  
 XX  
 PF 20-JAN-1989; 89CA-00588749.  
 XX  
 PR 20-JAN-1989; 89CA-00588749.  
 XX  
 PA (GEHO ) GEN HOSPITAL CORP.  
 XX  
 PI Seed B;  
 XX  
 DR WPI; 2000-063015/06.  
 DR P-PSDB; AAY59170.  
 XX  
 XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 XX Example 1; Page 47-53; 89pp; English.  
 XX  
 CC The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence encoding  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the DNA which  
 CC encodes the fusion protein CD4mm where the CD4 is linked to human IgG1  
 CC at the Mst2 site upstream of the CH1 region  
 XX  
 XX Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other:  
 SQ  
 Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGGCAACTGGGGCTCTCCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTAGGCACTGCTGTGGTGGCAACTGGGGCTCTCCCA 170  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACAGTGAATGACC 120  
 DB 171 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACAGTGAATGACC 230  
 QY 121 TGTACAGCTTCCCGAAGAGAGATCAATTCACCTGGAAAACTCCAAACAGATAAAG 180  
 DB 231 TGTACAGCTTCCCGAAGAGAGATCAATTCACCTGGAAAACTCCAAACAGATAAAG 290  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACCTAAAGGTCCATCAAGCTGAATGATGGGCT 240  
 DB 291 ATTCTGGGAATCAGGGCTCTCTTAACCTAAAGGTCCATCAAGCTGAATGATGGGCT 350  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 351 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 410  
 QY 301 ATAGAAGACTCAGATTAATCTGATCTGTGAAGTGAAGACCAAGAGAGAGAGTGCATTG 360  
 DB 411 ATAGAAGACTCAGATTAATCTGATCTGTGAAGTGAAGACCAAGAGAGAGAGTGCATTG 470

QY 361 CTAGTGTGGATTTGACTGCCAACTCTGACACCCACTCTGTCAGGGGACAGACCTGACC 420  
 DB 471 CTAGTGTGGATTTGACTGCCAACTCTGACACCCACTCTGTCAGGGGACAGACCTGACC 530  
 QY 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTGATGTCAAATGATGAGTCCAAAGGGT 480  
 DB 531 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTGATGTCAAATGATGAGTCCAAAGGGT 590  
 QY 481 AAAACATACAGGGGGGGAAGAACCTCTCGTGTCTCAGCTGGAGCTCCAGATTAAGGC 540  
 DB 591 AAAACATACAGGGGGGGAAGAACCTCTCGTGTCTCAGCTGGAGCTCCAGATTAAGGC 650  
 QY 541 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGTGGAGTTCAAAATAGACATGCTG 600  
 DB 651 ACCTGACATGCACTGTCTTGAGAAACCAAGAAAGTGGAGTTCAAAATAGACATGCTG 710  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736  
 RESULT 11  
 AAA35205  
 ID AAA35205 standard; DNA; 1742 BP.  
 XX  
 AC AAA35205;  
 XX  
 DT 28-JUL-2000 (first entry)  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:79.  
 XX  
 KM Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KM phosphorothioate; impaired respiration; inflammation; allergy;  
 KM allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KM antiallergic; antiasthmatic; cyostatic; analgesic; impaired airway;  
 KM lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KM respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KM pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KM cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 PF 03-AUG-1999; 99MO-US017712.  
 XX  
 PR 03-AUG-1998; 98US-0095212P.  
 XX  
 PA (UYEC-) UNIV EAST CAROLINA.  
 XX  
 PI Nyce JW;  
 XX  
 DR WPI; 2000-205971/18.  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers.  
 XX  
 PS Disclosure; Page 1245; 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an antisense  
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets  
 CC nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC antiasthmatic, cyostatic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,  
 CC impaired respiration, respiratory distress syndrome, pain, cystic

CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive  
 CC pulmonary disease (COPD), and cancers such as leukaemia, lymphomas,  
 CC carcinomas, and cancers which may metastasize to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of the  
 CC ONs reduces side effects. The A-containing ONs break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing  
 CC bronchoconstriction and inflammation. AAA3233 to AAA3512 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185  
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ  
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3233 to  
 CC AAA33992) are specifically claimed ONs from the present invention. N.B.  
 CC Sequences given in the disclosure of the present invention do not match  
 CC up with their corresponding SEQ ID NO: sequences given in the sequence  
 CC listing  
 CC XX  
 SQ Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGACCGGGGAGTCCCTTTTAAAGCACTTGCTGCTGCAACTGGGCTCTCCCA 60  
 Db 76 ATGACCGGGGAGTCCCTTTTAAAGCACTTGCTGCTGCTGCAACTGGGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGCTGCTGCGCAAAAAGGGGATACAGTGAATGACC 120  
 Db 136 GCAGCCACTCAGGAGAAAGAGTGCTGCTGCGCAAAAAGGGGATACAGTGAATGACC 195  
 QY 121 TGTACAGCTTCCAGAAAGAGCATTCATCTGGAATACTCCAAACCGATTAAG 180  
 Db 196 TGTACAGCTTCCAGAAAGAGCATTCATCTGGAATACTCCAAACCGATTAAG 255  
 QY 181 ATTCTGGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 240  
 Db 256 ATTCTGGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300  
 Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 375  
 QY 301 ATGAGAACTCAATTAATCTTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 360  
 Db 376 ATGAGAACTCAATTAATCTTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 435  
 QY 361 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 420  
 Db 436 CTAGTGTTCGGATTGACTGCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 495  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480  
 Db 496 CTGACCTTGGAGAGCCCTCTGTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 555  
 QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCTGCTGCTCACTGGAAGCTCCAGATAGTGC 540  
 Db 556 AAAAACAATACAGGGGGGAAAGACCTCTCTGCTGCTCACTGGAAGCTCCAGATAGTGC 615  
 QY 541 ACCTGACATGACCTGCTTGGAGAACCAAGAGGTGAGTCAAAATAGACATGTG 600  
 Db 616 ACCTGACATGACCTGCTTGGAGAACCAAGAGGTGAGTCAAAATAGACATGTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 Db 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 12  
 AAF21327

ID AAF21327 standard; DNA; 1742 BP.

XX AAF21327;  
 XX

DT 14-MAR-2001 (first entry)  
 XX Human low adenosine antisense oligonucleotide related sequence #2894.  
 XX  
 XX Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 KW human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiasthmatic; analgesic; hypotensive; cytostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.  
 XX  
 XX Homo sapiens.  
 OS  
 XX  
 PN W0200062736-A2.  
 PD 26-OCT-2000.  
 PF 24-MAR-2000; 2000WO-US008020.  
 PR 06-APR-1999; 99US-0127958P.  
 PA (UYEC-) UNIV EAST CAROLINA.  
 PY (NYCE/) NYCE J W.  
 PI Nycy JW;  
 XX  
 XX WPI; 2000-679539/66.  
 XX  
 XX Low adenosine (A) content antisense oligonucleotides which do not trigger  
 PT adenosine receptors during metabolism, useful e.g. for treating cancers  
 PT and respiratory obstructions.  
 PT  
 PS Disclosure, Page 1329; 1592pp; English.  
 XX  
 XX The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiasthmatic, hypotensive and cytostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and/or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulin and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokine and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensins, growth factors, vasoactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction  
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or  
 CC surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impeded respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention  
 CC XX  
 SQ Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGCGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGAGGAAAGAAAGTGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
 DB 136 GCAGCCACTCAGAGGAAAGAAAGTGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 195  
 QY 121 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCATTAAG 180  
 DB 196 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCATTAAG 255  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGCTCATCAGTGAATGATCGGCT 240  
 DB 256 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGCTCATCAGTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 375  
 QY 301 ATGAAGACTCAGTACTTACTTACTGTGGAAGTGAAGCAAGAGAGAGTGCATTG 360  
 DB 376 ATGAAGACTCAGTACTTACTTACTGTGGAAGTGAAGCAAGAGAGAGTGCATTG 435  
 QY 361 CTAGTGTTCGATGAGTGCCTCACTGACCACTGACCACTGCTTCAGGGGACAGGCTGACC 420  
 DB 436 CTAGTGTTCGATGAGTGCCTCACTGACCACTGACCACTGCTTCAGGGGACAGGCTGACC 495  
 QY 421 CTGACCTTGGAGAGACCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 DB 496 CTGACCTTGGAGAGACCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 555  
 QY 481 AAAAACTACAGGGGGGAGAGACCTCTCTGTCTCAGTGCATGTAGAGTCCAAAGGGGT 540  
 DB 556 AAAAACTACAGGGGGGAGAGACCTCTCTGTCTCAGTGCATGTAGAGTCCAAAGGGGT 615  
 QY 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
 DB 616 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701  
 RESULT 13  
 ID AAI65462 standard; DNA; 1742 BP.  
 AC AAI65462;  
 XX 10-DEC-2001 (first entry)  
 DT Nucleotide sequence of a human polynucleotide.  
 XX Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;  
 KM C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.  
 OS Homo sapiens.  
 PN W0200164752-A2.  
 PD 07-SEP-2001.  
 XX 28-FEB-2001; 2001WO-US006322.  
 PF 02-MAR-2000; 2000US-00517605.  
 PR (UTNY ) UNIV NEW YORK STATE.  
 XX (UTNY-) UNIV NIMEGEN.  
 PA Littman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;

XX WPI; 2001-602565/68.  
 DR An antibody for the treatment or prevention of HIV-infection comprises a  
 XX gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
 PT DC-SIGN due to concomitant conformational change.  
 PT Disclosure; Page 121-122; 131pp; English.  
 XX The specification describes an antibody which is specific for an  
 CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
 CC is exposed upon gp120 binding of DC-SIGN due to concomitant  
 CC conformational change. DC-SIGN is a receptor that is specifically  
 CC expressed on dendritic cells and facilitates infection of T lymphocytes  
 CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
 CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
 CC with high affinity. The antibody of the invention inhibits the trans  
 CC enhancement of HIV entry into a T cell or macrophage facilitated by  
 CC dendritic cells. The antibody is useful to treat or prevent HIV  
 CC infection. The present sequence represents a human polynucleotide, which  
 CC is used in the course of the invention  
 XX Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;  
 SQ  
 Query Match 43.4%; Score 624.4; DB 4; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1,8e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTTGTGCTGCAACTGCGCTCTCCCA 135  
 QY 61 GCAGCCACTCAGAGGAAAGAAAGTGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
 DB 136 GCAGCCACTCAGAGGAAAGAAAGTGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 195  
 QY 121 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCATTAAG 180  
 DB 196 TGTACACTTCCCAAGAAAGATACATTCATCTGGAAGAACTCCACCATTAAG 255  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGCTCATCAGTGAATGATCGGCT 240  
 DB 256 ATTCTGGAAATCAGGGCTCTCTTCTTAACCTAAGCTCATCAGTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
 DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 375  
 QY 301 ATGAAGACTCAGTACTTACTTACTGTGGAAGTGAAGCAAGAGAGAGTGCATTG 360  
 DB 376 ATGAAGACTCAGTACTTACTTACTGTGGAAGTGAAGCAAGAGAGAGTGCATTG 435  
 QY 361 CTAGTGTTCGATGAGTGCCTCACTGACCACTGACCACTGCTTCAGGGGACAGGCTGACC 420  
 DB 436 CTAGTGTTCGATGAGTGCCTCACTGACCACTGACCACTGCTTCAGGGGACAGGCTGACC 495  
 QY 421 CTGACCTTGGAGAGACCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 480  
 DB 496 CTGACCTTGGAGAGACCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGGT 555  
 QY 481 AAAAACTACAGGGGGGAGAGACCTCTCTGTCTCAGTGCATGTAGAGTCCAAAGGGGT 540  
 DB 556 AAAAACTACAGGGGGGAGAGACCTCTCTGTCTCAGTGCATGTAGAGTCCAAAGGGGT 615  
 QY 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
 DB 616 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTCAAAATAGACATGCTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701



PT altered expression of genes in an immunological response or for  
PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma  
PT or osteoarthritis.

XX  
PS Claim 1, SEQ ID NO 1013, 50bp, English.

CC The invention relates to a composition comprising a plurality of cDNAs  
CC for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperosinophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNA, cDNAs or  
CC genomic fragments. This sequence represents a human cDNA of the  
CC invention. Note: The sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at seqdata.uspto.gov/sequence.html.

XX  
SQ Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 11; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.8e-129;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCAACTGGCGCTCTCCCA 60  
Db 76 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGTGCAACTGGCGCTCTCCCA 135  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAAC 120  
Db 136 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAAC 195  
QY 121 TGTACAGCTTCCAGAGAAGACATACATTCACCTGGAATAAATCCCAACGATAAAG 180  
Db 196 TGTACAGCTTCCAGAGAAGACATACATTCACCTGGAATAAATCCCAACGATAAAG 255  
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTCATCAAGCTGAATGCGCT 240  
Db 256 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTCATCAAGCTGAATGCGCT 315  
QY 241 GACTCAAGAGAAGCCTTTGGGACCAAGAAATTTCCCTGATCATTAAGATCTTAAG 300  
Db 316 GACTCAAGAGAAGCCTTTGGGACCAAGAAATTTCCCTGATCATTAAGATCTTAAG 375  
QY 301 ATAGAAGACTCAGATTAATCATCTGTAAGTGAGAGACAGAAAGAGAGAGTGAATTG 360  
Db 376 ATAGAAGACTCAGATTAATCATCTGTAAGTGAGAGACAGAAAGAGAGAGTGAATTG 435  
QY 361 CTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 420  
Db 436 CTAGTGTTCGATTAAGTCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGCAATGTAGAGATCAAGGGGT 480  
Db 496 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGCAATGTAGAGATCAAGGGGT 555  
QY 481 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTAGCTGAGCTCCAGATATGTGC 540

Db 556 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTAGCTGAGCTCCAGATATGTGC 615  
QY 541 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATGCTG 600  
Db 616 ACCTGACATGCACTGTCTTGGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATGCTG 675  
QY 601 GTGCTAGCTTCCAGAAAGGCTTCGG 626  
Db 676 GTGCTAGCTTCCAGAAAGGCTTCAG 701

Search completed: September 20, 2005, 10:37:09  
Job time : 870 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 20, 2005, 08:29:41 ; Search time 298 Seconds  
(without alignments)  
7906.841 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440  
Sequence: 1 atgaaccgggagctccctt.....gtccagcagcacaataactcg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapect 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: Issued\_Patents\_NA:\*  
2: /cgn2\_6/ptodata/1/ina/5A COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/5B COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6A COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/6B COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfile1.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	43.4	1213	6 5223418-1	Patent No. 5223418
2	625	43.4	1213	6 5223418-1	Patent No. 5223418
3	624.4	43.4	1415	5 PCT-US92-08090-4	Sequence 4, Appl1
4	624.4	43.4	1421	5 PCT-US92-08090-3	Sequence 3, Appl1
5	624.4	43.4	1448	5 PCT-US92-08090-2	Sequence 2, Appl1
6	624.4	43.4	1742	3 US-09-517-605-7	Sequence 7, Appl1
7	624.4	43.4	1742	3 US-09-023-655-1013	Sequence 1013, Ap
8	624.4	43.4	2465	5 PCT-US92-08090-1	Sequence 1, Appl1
9	622.8	43.2	1304	2 US-08-284-391B-28	Sequence 28, Appl1
10	622.8	43.2	1304	3 US-09-218-950-28	Sequence 28, Appl1
11	622.8	43.2	1304	4 US-08-394-388A-28	Sequence 28, Appl1
12	622.8	43.2	1389	2 US-08-417-495-2	Sequence 2, Appl1
13	622.8	43.2	1389	2 US-08-284-391B-2	Sequence 2, Appl1
14	622.8	43.2	1389	3 US-09-218-950-2	Sequence 2, Appl1
15	622.8	43.2	1389	4 US-08-394-388A-2	Sequence 2, Appl1
16	622.8	43.2	1389	5 PCT-US92-01785-2	Sequence 2, Appl1
17	622.8	43.2	1389	5 PCT-US95-00454-2	Sequence 2, Appl1
18	622.8	43.2	1599	2 US-08-417-495-3	Sequence 3, Appl1
19	622.8	43.2	1599	2 US-08-284-391B-3	Sequence 3, Appl1
20	622.8	43.2	1599	3 US-09-218-950-3	Sequence 3, Appl1
21	622.8	43.2	1599	4 US-08-394-388A-3	Sequence 3, Appl1
22	622.8	43.2	1599	5 PCT-US92-01785-3	Sequence 3, Appl1
23	622.8	43.2	1599	5 PCT-US95-00454-3	Sequence 3, Appl1
24	622.8	43.2	1728	2 US-08-417-495-1	Sequence 1, Appl1
25	622.8	43.2	1728	2 US-08-284-391B-1	Sequence 1, Appl1
26	622.8	43.2	1728	3 US-09-218-950-1	Sequence 1, Appl1
27	622.8	43.2	1728	4 US-08-394-388A-1	Sequence 1, Appl1

28	622.8	43.2	1728	5 PCT-US92-01785-1	Sequence 1, Appl1
29	622.8	43.2	1728	5 PCT-US95-00454-1	Sequence 1, Appl1
30	622.8	43.2	1742	3 US-08-466-368-3	Sequence 3, Appl1
31	622.8	43.2	1742	4 US-08-470-998-1	Sequence 1, Appl1
32	622.8	43.2	1742	4 US-08-328-500-8	Sequence 8, Appl1
33	622.8	43.2	2589	4 US-08-472-888A-5	Sequence 5, Appl1
34	621.2	43.1	1273	3 US-08-466-368-1	Sequence 1, Appl1
35	621.2	43.1	1273	4 US-08-328-500-1	Sequence 1, Appl1
36	621.2	43.1	1416	3 US-08-236-311-2	Sequence 2, Appl1
37	621.2	43.1	1416	3 US-08-457-918-2	Sequence 2, Appl1
38	621.2	43.1	1416	4 US-10-157-408-2	Sequence 2, Appl1
39	621.2	43.1	2940	2 US-08-418-848A-3	Sequence 3, Appl1
40	621.2	43.1	3426	2 US-08-418-848A-2	Sequence 2, Appl1
41	621.2	43.1	3480	2 US-08-418-848A-4	Sequence 4, Appl1
42	621.2	43.1	3721	2 US-08-418-848A-5	Sequence 5, Appl1
43	621.2	43.1	3993	2 US-08-418-848A-6	Sequence 6, Appl1
44	621.2	43.1	4059	2 US-08-418-848A-7	Sequence 7, Appl1
45	621.2	43.1	4632	2 US-08-418-848A-8	Sequence 8, Appl1

ALIGNMENTS

RESULT 1  
5223418-1  
; Patent No. 5223418  
; APPLICANT: ARCURI, EDWARD J.;BRAUNER, MARY E.; DONOVAN, MARY  
; J.; GERBER, ROBERT G.; KELLER, JOHN A.  
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF  
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS  
; NUMBER OF SEQUENCES: 2  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/589,979  
; FILING DATE: 28-SEP-1990  
; SEQ ID NO:1:  
; LENGTH: 1213  
5223418-1

Query Match 43.4%; Score 625; DB 6; Length 1213;  
Best Local Similarity 97.7%; Pred. No. 3.7e-138;  
Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY	1	ATGAACGGGAGATCCCTTTTAAAGCACTTGCTGTGCTGCACTGGCGCTCTCCCA	60
DB	79	ATGAACGGGAGATCCCTTTTAAAGCACTTGCTGTGCTGCACTGGCGCTCTCCCA	138
QY	61	GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC	120
DB	139	GCAGCCACTCAGGAGAAAGTGTGCTGGGCAAAAAGGGATACAGTGAACCTGACC	198
QY	121	TGTACAGCTTCCAGAAAGCATCAATTCATGGAATAAATCCCAACGATTAAG	180
DB	199	TGTACAGCTTCCAGAAAGCATCAATTCATGGAATAAATCCCAACGATTAAG	258
QY	181	ATTCTGGGAATCAGGGCTCTCTTAATCAAAAGTCAATCCAAAGTGAATGCGCT	240
DB	259	ATTCTGGGAATCAGGGCTCTCTTAATCAAAAGTCAATCCAAAGTGAATGCGCT	318
QY	241	GACTCAAGAAAGCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGATTTAAG	300
DB	319	GACTCAAGAAAGCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGATTTAAG	378
QY	301	ATGAAGAGCTCAGATCTTAACATCTGTGAAGTGAAGGACCAAGAAAGGAGTGAATG	360
DB	379	ATGAAGAGCTCAGATCTTAACATCTGTGAAGTGAAGGACCAAGAAAGGAGTGAATG	438
QY	361	CTAGTGTGAGATGATGCTGCAACTGCAACCTGCTTCAGGGGAGAGCTGACC	420
DB	439	CTAGTGTGAGATGATGCTGCAACTGCAACCTGCTTCAGGGGAGAGCTGACC	498
QY	421	CTGACCTGGAGAGCCCCCTGTGATAGCCCTTCAGTGCATATGAGATCCAAAGGGT	480
DB	499	CTGACCTGGAGAGCCCCCTGTGATAGCCCTTCAGTGCATATGAGATCCAAAGGGT	558

```

QY      481  AAAAATACAGAGGGGGAAGACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 540
        |||
Db      559  AAAAATACAGAGGGGGAAGACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 618

QY      541  ACCTGACATGCACTGCTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600
        |||
Db      619  ACCTGACATGCACTGCTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 678

QY      601  GTGCTAGCTTCCAGAGGCTCCGAGAGTGCGGCTAGTGGGGAAGCG 649
        |||
Db      679  GTGCTAGCTTCCAGAGGCTCCGAGATGCTATAGAAAGAGGCG 727

```

```

RESULT 2
5223418-1
; Patent No. 5223418
; APPLICANT: ARCURI, EDWARD J.; BRANNER, MARY E.; DONOVAN, MARY
; J.; GERBER, ROBERT G.; KELLER, JOHN A.
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS
; NUMBER OF SEQUENCES: 2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/07/589,979
; FILING DATE: 28-SEP-1990
; SEQ ID NO:1:
; LENGTH: 1213
5223418-1

```

```

Query Match      43.4%; Score 625; DB 6; Length 1213;
Best Local Similarity 97.7%; Pred. No. 3,7e-138;
Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY      1  ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 60
        |||
Db      79  ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 138

QY      61  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
        |||
Db      139  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 198

QY      121  TGTACAGCTTCCAGAGAAAGAGATCAATTCACCTGGAAGAAATCCACAGATTAAG 180
        |||
Db      199  TGTACAGCTTCCAGAGAAAGAGATCAATTCACCTGGAAGAAATCCACAGATTAAG 258

QY      181  ATTCTGGGAATCAGGGCTCTCTTAATAAGTCCATCAAGCTGAATGATCGGCT 240
        |||
Db      259  ATTCTGGGAATCAGGGCTCTCTTAATAAGTCCATCAAGCTGAATGATCGGCT 318

QY      241  GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300
        |||
Db      319  GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 378

QY      301  ATGAAGACTCAATATCTTACATCTGGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
        |||
Db      379  ATGAAGACTCAATATCTTACATCTGGAAGTGAAGACCAAGAAAGAGGTGCAATTG 438

QY      361  CTAGTGTTCGATTTGACTGCAACTGACACCACTGCTTCAAGGGGCAAGCTGACC 420
        |||
Db      439  CTAGTGTTCGATTTGACTGCAACTGACACCACTGCTTCAAGGGGCAAGCTGACC 498

QY      421  CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 480
        |||
Db      499  CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAAAGGGT 558

QY      481  AAAAATACAGAGGGGGAAGACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 540
        |||
Db      559  AAAAATACAGAGGGGGAAGACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 618

QY      541  ACCTGACATGCACTGCTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600
        |||
Db      619  ACCTGACATGCACTGCTCTTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 678

```

```

QY      601  GTGCTAGCTTTCAGAGGCTTCCGAGAGTGCGGCTAGTGGGGAAGCG 649
        |||
Db      679  GTGCTAGCTTTCAGAGGCTTCCAGATAGTCTATAGAAAGAGGCG 727

```

```

RESULT 3
PCT-US92-08090-4
; Sequence 4, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: AIDS
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kipatrick & Cody
; STREET: 100 Peachtree Street
; City: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patricia L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMR129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1415 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1410..1415
; OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-4

```

```

Query Match      43.4%; Score 624.4; DB 5; Length 1415;
Best Local Similarity 99.8%; Pred. No. 5.4e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1  ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 60
        |||
Db      81  ATGAACCGGGGAGTCCCTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 140

QY      61  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
        |||
Db      141  GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 200

```

QY 121 TGTACAGCTTCCGAGAGAGATCAATTCCTGAGAAAACTCCAAACAGATTAAG 180  
 Db 201 TGTACAGCTTCCGAGAGAGATCAATTCCTGAGAAAACTCCAAACAGATTAAG 260  
 QY 181 ATTCTGGGAAATCAGGGCTCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 240  
 Db 261 ATTCTGGGAAATCAGGGCTCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 320  
 QY 241 GACTCAGAGAAAGGCTTTTGGGACCAAGAACTTCCCTGATCAATCAAGAACTTTAAG 300  
 Db 321 GACTCAGAGAAAGGCTTTTGGGACCAAGAACTTCCCTGATCAATCAAGAACTTTAAG 380  
 QY 301 ATAGAAAGCTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
 Db 381 ATAGAAAGCTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 440  
 QY 361 CTAGGTGTTGGATTGACTGCTCAACTGACACCCAGCTGCTTCAGGGGACAGAGCTGACC 420  
 Db 441 CTAGGTGTTGGATTGACTGCTCAACTGACACCCAGCTGCTTCAGGGGACAGAGCTGACC 500  
 QY 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGAATGATGATCCAAAGGGGT 480  
 Db 501 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGAATGATGATCCAAAGGGGT 560  
 QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGC 540  
 Db 561 AAAAACAATACAGGGGGGAGAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGACATGACATGCTTGTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 600  
 Db 621 ACCTGACATGACATGCTTGTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 680  
 QY 601 GTGCTAGCTTCCAGAGAGGCTCCGG 626  
 Db 681 GTGCTAGCTTCCAGAGAGGCTCCAG 706

## RESULT 4

PCT-US92-08090-3  
 ; Sequence 3, Application PC/TUS9208090  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Tang, J. N.  
 ; TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
 ; TITLE OF INVENTION: Aids  
 ; NUMBER OF SEQUENCES: 15  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: Kilpatrick & Cody  
 ; STREET: 100 Peachtree Street  
 ; CITY: Atlanta  
 ; STATE: Georgia  
 ; COUNTRY: U.S.  
 ; ZIP: 30303  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: PatentIn Release #1.0, Version #1.25  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/US92/08090  
 ; FILING DATE: 19920922  
 ; CLASSIFICATION:  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Pabst, Patricia L.  
 ; REGISTRATION NUMBER: 31,284  
 ; REFERENCE/DOCKET NUMBER: OMR#129  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 404-572-6508  
 ; TELEFAX: 404-572-6555  
 ; INFORMATION FOR SEQ ID NO: 3:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1421 base pairs  
 ; TYPE: NUCLEIC ACID

; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: DNA (genomic)  
 ; HYPOETHERICAL: NO  
 ; ANTI-SENSE: NO  
 ; ORIGINAL SOURCE:  
 ; ORGANISM: Homo sapien  
 ; TISSUE TYPE: Epithelial  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1..6  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; FEATURE:  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1275..1280  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; NAME/KEY: misc\_feature  
 ; LOCATION: 1416..1421  
 ; OTHER INFORMATION: /note= "Restriction site"  
 ; PCT-US92-08090-3

Query Match 43.4%; Score 624.4; DB 5; Length 1421;  
 Best Local Similarity 99.8%; Pred. No. 5,4e-138;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGATCCCTTTTATGACATGCTTCTGTGCTGCAACTGGCGCTCCCA 60  
 Db 81 ATGACCGGGAGATCCCTTTTATGACATGCTTCTGTGCTGCAACTGGCGCTCCCA 140  
 QY 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATA CAGTGAATGACC 120  
 Db 141 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGGATA CAGTGAATGACC 200  
 QY 121 TGTACAGCTTCCGAGAGAGATCAATTCCTGAGAAAACTCCAAACAGATTAAG 180  
 Db 201 TGTACAGCTTCCGAGAGAGATCAATTCCTGAGAAAACTCCAAACAGATTAAG 260  
 QY 181 ATTCTGGGAAATCAGGGCTCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 240  
 Db 261 ATTCTGGGAAATCAGGGCTCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 320  
 QY 241 GACTCAGAGAAAGGCTTTTGGGACCAAGAACTTCCCTGATCAATCAAGAACTTTAAG 300  
 Db 321 GACTCAGAGAAAGGCTTTTGGGACCAAGAACTTCCCTGATCAATCAAGAACTTTAAG 380  
 QY 301 ATAGAAAGCTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGTCAATTG 360  
 Db 381 ATAGAAAGCTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGTCAATTG 440  
 QY 361 CTAGGTGTTGGATTGACTGCTCAACTGACACCCAGCTGCTTCAGGGGACAGAGCTGACC 420  
 Db 441 CTAGGTGTTGGATTGACTGCTCAACTGACACCCAGCTGCTTCAGGGGACAGAGCTGACC 500  
 QY 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGAATGATGATCCAAAGGGGT 480  
 Db 501 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGAATGATGATCCAAAGGGGT 560  
 QY 481 AAAAACAATACAGGGGGGAGAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGC 540  
 Db 561 AAAAACAATACAGGGGGGAGAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATATGAGC 620  
 QY 541 ACCTGACATGACATGCTTGTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 600  
 Db 621 ACCTGACATGACATGCTTGTGACAGAACCAAGAAAGGTGAGATTCAAAATATGACATGCTG 680  
 QY 601 GTGCTAGCTTCCAGAGAGGCTCCGG 626  
 Db 681 GTGCTAGCTTCCAGAGAGGCTCCAG 706

## RESULT 5

PCT-US92-08090-2

```

; Sequence 2, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; TITLE OF INVENTION: Aids
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMFP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: Nucleic Acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1444..1448
; OTHER INFORMATION: /note= "Restriction site"
; PCT-US92-08090-2

Query Match 43.4%; Score 624.4; DB 5; Length 1448;
Best Local Similarity 99.8%; Pred. No. 5.4e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGGGCTCCCTCCA 60
DB 81 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGGGCTCCCTCCA 140
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 141 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 200
QY 121 TGTACAGCTTCCAGAAAGAAAGCATATTCCTACTGAAAAAAGTCCAAACGATTAAG 180
DB 201 TGTACAGCTTCCAGAAAGAAAGCATATTCCTACTGAAAAAAGTCCAAACGATTAAG 260
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATGATCGGCT 240

```

```

DB 261 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATATCGCGCT 320
QY 241 GACTCAAGAAAGAGCTTTGGAGCAAGAAATCTTCCCTGATCATCAAGATCTTAAG 300
DB 321 GACTCAAGAAAGAGCTTTGGAGCAAGAAATCTTCCCTGATCATCAAGATCTTAAG 380
QY 301 ATGAAAGACTCAGATATCTTAATCTGTGTAAGTGAAGACACAGAAAGAGAGGTGCAATTG 360
DB 381 ATGAAAGACTCAGATATCTTAATCTGTGTAAGTGAAGACACAGAAAGAGAGGTGCAATTG 440
QY 361 CTAGTGTGGATTGATGCTGCAACTGTGACACCACTGCTCAGGGGAGAGCTGACC 420
DB 441 CTAGTGTGGATTGATGCTGCAACTGTGACACCACTGCTCAGGGGAGAGCTGACC 500
QY 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTGATGCAATGTAAGAGTCCAAAGGCT 480
DB 501 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTGATGCAATGTAAGAGTCCAAAGGCT 560
QY 481 AAAACATACAGGGGGGGAAGACCTCTCGGTCTCAGCTGAGCTCCAGATATGTGCT 540
DB 561 AAAACATACAGGGGGGGAAGACCTCTCGGTCTCAGCTGAGCTCCAGATATGTGCT 620
QY 541 ACCTGACATGCACTGTCTTGCAAGACAGAAAGTGAAGTCAAAATAGACATCGTG 600
DB 621 ACCTGACATGCACTGTCTTGCAAGACAGAAAGTGAAGTCAAAATAGACATCGTG 680
QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626
DB 681 GTGCTAGCTTTCCAGAGGCTCCGG 706

RESULT 6
US-09-517-605-7
; Sequence 7, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Liltman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; TITLE OF INVENTION: CELLS
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517, 605
; CURRENT FILING DATE: 2000-03-02
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-517-605-7

Query Match 43.4%; Score 624.4; DB 3; Length 1742;
Best Local Similarity 99.8%; Pred. No. 5.7e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGGGCTCCCTCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGCTGCAACTGGGGCTCCCTCCA 135
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120
DB 136 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 195
QY 121 TGTACAGCTTCCAGAAAGAAAGCATATTCCTACTGAAAAAAGTCCAAACGATTAAG 180
DB 196 TGTACAGCTTCCAGAAAGAAAGCATATTCCTACTGAAAAAAGTCCAAACGATTAAG 255
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 256 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATGATCGGCT 315

```

```

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
| | | | |
DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
| | | | |
QY 301 ATGAAGACTCAGTACTTACTTACTTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
| | | | |
DB 376 ATGAAGACTCAGTACTTACTTACTTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 435
| | | | |
QY 361 CTGAGTTCGGATGACCTGCAACTGACACCCACCTGCTTCAAGGGGAGAGGCTGACC 420
| | | | |
DB 436 CTGAGTTCGGATGACCTGCAACTGACACCCACCTGCTTCAAGGGGAGAGGCTGACC 495
| | | | |
QY 421 CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAGGGGT 480
| | | | |
DB 496 CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAGGGGT 555
| | | | |
QY 481 AAAAAGCTACAGGGGGGAGAGACCTCTCTGCTGTCTAGCTGAGAGCTCCAGATAGTGGC 540
| | | | |
DB 556 AAAAAGCTACAGGGGGGAGAGACCTCTCTGCTGTCTAGCTGAGAGCTCCAGATAGTGGC 615
| | | | |
QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGGTGAGTCAAAATAGACATGCTG 600
| | | | |
DB 616 ACCTGACATGACCTGTCTTGCAGAACCAAGAGGTGAGTCAAAATAGACATGCTG 675
| | | | |
QY 601 GTGCTAGCTTTCCAGAGGCTCCCG 626
| | | | |
DB 676 GTGCTAGCTTTCCAGAGGCTCCCG 701
| | | | |

```

RESULT 7  
US-09-023-655-1013  
Sequence 1013, Application US/09023655

Patent No. 6607879

GENERAL INFORMATION:

APPLICANT: Cocks, Benjamin G.

APPLICANT: Susan G. Stuart

APPLICANT: Jeffrey J. Seilhamer

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE

NUMBER OF SEQUENCES: 1508

CORRESPONDENCE ADDRESS:

ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

STREET: 3174 PORTER DRIVE

CITY: PALO ALTO

STATE: CALIFORNIA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/023,655

FILING DATE: HEREMITH

CLASSIFICATION:

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Zeller, Karen J.

REGISTRATION NUMBER: 37,071

REFERENCE/DOCKET NUMBER: PA-0001 US

TELEPHONE: (650) 855-0555

TELEFAX: (650) 845-4166

INFORMATION FOR SEQ ID NO: 1013:

SEQUENCE CHARACTERISTICS:

LENGTH: 1742 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: G179143  
US-09-023-655-1013

Query Match 43.4%; Score 624.4; DB 4; Length 1742;

Best Local Similarity 99.8%; Pred. No. 5,7e-138; Indels 0; Gaps 0;

Matches 625; Conservative 0; Mismatches 1;

```

QY 1 ATGAACGGGGAGTCCCTTTTGAAGCACTGCTCTGTGTGTGCACTGGGCGCTCCCA 60
| | | | |
DB 76 ATGAACGGGGAGTCCCTTTTGAAGCACTGCTCTGTGTGTGCACTGGGCGCTCCCA 135
| | | | |
QY 61 GCAGCACTCAGGAAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
| | | | |
DB 136 GCAGCACTCAGGAAAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 195
| | | | |
QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCAGTGGAAAACTCCAACTGATTAAG 180
| | | | |
DB 196 TGTACAGCTTCCAGAAAGAGCATTCATTCAGTGGAAAACTCCAACTGATTAAG 255
| | | | |
QY 181 ATTCTGGAAATCAGGGCTCCTTTAATCAATGATCAATCCAGCTGAATGCGGCT 240
| | | | |
DB 256 ATTCTGGAAATCAGGGCTCCTTTAATCAATGATCAATCCAGCTGAATGCGGCT 315
| | | | |
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
| | | | |
DB 316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
| | | | |
QY 301 ATGAAGACTCAGTACTTACTTACTTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360
| | | | |
DB 376 ATGAAGACTCAGTACTTACTTACTTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 435
| | | | |
QY 361 CTGAGTTCGGATGACCTGCAACTGACACCCACCTGCTTCAAGGGGAGAGGCTGACC 420
| | | | |
DB 436 CTGAGTTCGGATGACCTGCAACTGACACCCACCTGCTTCAAGGGGAGAGGCTGACC 495
| | | | |
QY 421 CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAGGGGT 480
| | | | |
DB 496 CTGACCTTGAAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGATCCAGGGGT 555
| | | | |
QY 481 AAAAAGCTACAGGGGGGAGAGACCTCTCTGCTGTCTAGCTGAGAGCTCCAGATAGTGGC 540
| | | | |
DB 556 AAAAAGCTACAGGGGGGAGAGACCTCTCTGCTGTCTAGCTGAGAGCTCCAGATAGTGGC 615
| | | | |
QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGGTGAGTCAAAATAGACATGCTG 600
| | | | |
DB 616 ACCTGACATGACCTGTCTTGCAGAACCAAGAGGTGAGTCAAAATAGACATGCTG 675
| | | | |
QY 601 GTGCTAGCTTTCCAGAGGCTCCCG 626
| | | | |
DB 676 GTGCTAGCTTTCCAGAGGCTCCCG 701
| | | | |

```

RESULT 8  
PCT-US92-08090-1  
Sequence 1, Application PC/TUS9208090

GENERAL INFORMATION:

APPLICANT: Tang, J. N.

TITLE OF INVENTION: Fusion Protein Genes for Treatment of

TITLE OF INVENTION: Aids

NUMBER OF SEQUENCES: 15

CORRESPONDENCE ADDRESS:

ADDRESSEE: Kilpatrick & Cody

STREET: 100 Peachtree Street

CITY: Atlanta

STATE: Georgia

COUNTRY: U.S.

ZIP: 30303

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

```

SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08090
FILING DATE: 19920922
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRF129
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2465 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Epithelial
FEATURE:
NAME/KEY: misc_feature
LOCATION: 1..6
OTHER INFORMATION: /note= "Restriction site"
NAME/KEY: misc_feature
LOCATION: 2460..2465
OTHER INFORMATION: /note= "Restriction site"
PCT-US92-08090-1

```

```

Query Match      43.4%; Score 624.4; DB 5; Length 2465;
Best Local Similarity 99.8%; Pred. No. 6.3e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60
DB 81 ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 140
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATACC 120
DB 141 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATACC 200
QY 121 TGTACAGCTTCCGGAAGAGACATCAATTCACATGGAATACTCAACCAAGTAAAG 180
DB 201 TGTACAGCTTCCGGAAGAGACATCAATTCACATGGAATACTCAACCAAGTAAAG 260
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 261 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 320
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 321 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 380
QY 301 ATGAAGACTCAGATTACTTACATCTGGAAGTGAAGACCAAGAAAGGAGGTCAATTG 360
DB 381 ATGAAGACTCAGATTACTTACATCTGGAAGTGAAGACCAAGAAAGGAGGTCAATTG 440
QY 361 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAGGGGCAAGAGCTGACC 420
DB 441 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAGGGGCAAGAGCTGACC 500
QY 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480
DB 501 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 560
QY 481 AAAAACAATACAGGGGGGGAAGAACCTCTCCGCTGCTAGCTGAGAGTCCAGGATATGAGC 540
DB 561 AAAAACAATACAGGGGGGGAAGAACCTCTCCGCTGCTAGCTGAGAGTCCAGGATATGAGC 620

```

```

QY 541 ACCTGACATGCACTGTCTTGAAGAACCAAGAAAGTGGAGTTCAAAATAGACATCTG 600
DB 621 ACCTGACATGCACTGTCTTGAAGAACCAAGAAAGTGGAGTTCAAAATAGACATCTG 680
QY 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626
DB 681 GTGCTAGCTTCCAGAAAGGCTCCAG 706

```

## RESULT 9

```

US-08-284-391B-28
Sequence 28, Application US/08284391B
Patent No. 5851828
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSER: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1304 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-284-391B-28

```

```

Query Match      43.2%; Score 622.8; DB 2; Length 1304;
Best Local Similarity 99.7%; Pred. No. 1.3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60
DB 111 ATGAACCGGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 170
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATGAGC 120
DB 171 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATGAGC 230

```



```

APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESS: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/394,388A
FILING DATE: 24-FEB-1995
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 28:
SEQUENCE CHARACTERISTICS:
LENGTH: 1304 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-08-394-388A-28

```

```

Query Match 43.2%; Score 622.8; DB 4; Length 1304;
Best Local Similarity 99.7%; Pred. No. 1.3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGAACGGGAGGAGCCCTTTAGGCACTGCTCTGCTGCAAGCGGGCTCCCTCCA 60
DB 111 ATGAACGGGAGGAGCCCTTTAGGCACTGCTCTGCTGCAAGCGGGCTCCCTCCA 170
QY 61 GCAGCCACTCAGGAGAAAGAGTGTCTGTGGCAAAAAAGGGGATACAGTGAAGTACC 120
DB 171 GCAGCCACTCAGGAGAAAGAGTGTGTCTGTGGCAAAAAAGGGGATACAGTGAAGTACC 230
QY 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCCACTGGAATACTCAACAGATTAAG 180
DB 231 TGTACAGCTTCCCAAGAAAGAGATCAATTCCACTGGAATACTCAACAGATTAAG 290
QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAAAGTGAATGCGGCT 240
DB 291 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCCAAAGTGAATGCGGCT 350
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTAAG 300
DB 351 GACTCAAGAAAGCCTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTAAG 410

```

```

QY 301 ATGAAAGACTCAGATCTTAACTTGTGAATGAGAGACCAAGAGAGAGGTGCATTTG 360
DB 411 ATGAAAGACTCAGATCTTAACTTGTGAATGAGAGACCAAGAGAGAGGTGCATTTG 470
QY 361 CTAGTGTGGATTGATCTGCAACTCTGACCCACCTGCTTCAAGGGGAGAGCTGACC 420
DB 471 CTAGTGTGGATTGATCTGCAACTCTGACCCACCTGCTTCAAGGGGAGAGCTGACC 530
QY 421 CTGACCTTGGAGAGCCCTGATAGTACCCCTCAGTGCATATGATAGAGTCCAAAGGCT 480
DB 531 CTGACCTTGGAGAGCCCTGATAGTACCCCTCAGTGCATATGATAGAGTCCAAAGGCT 590
QY 481 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTCACTGAGAGTCCAGATATGAGC 540
DB 591 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTCACTGAGAGTCCAGATATGAGC 650
QY 541 ACCTGACATGACCTGCTTGGAGAACCAAGAGGTGAGATTTAAATATAGCATCTG 600
DB 651 ACCTGACATGACCTGCTTGGAGAACCAAGAGGTGAGATTTAAATATAGCATCTG 710
QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626
DB 711 GTGCTAGCTTTCCAGAGGCTCCGAG 736

```

```

RESULT 12
US-08-417-495-2
Sequence 2, Application US/08417495
Patent No. 5843728
GENERAL INFORMATION:
APPLICANT: Seed, Brian et al.
TITLE OF INVENTION: Redirection of Cellular Immunity by
NUMBER OF SEQUENCES: 27
CORRESPONDENCE ADDRESS:
ADDRESS: Fish & Richardson
STREET: 225 Franklin Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110-2804
COMPUTER READABLE FORM:
MEDIUM TYPE: 3.5" Diskette, 1.44 MB
COMPUTER: IBM PS/2 Model 50Z or 55SX
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
SOFTWARE: Wordperfect (Version 5.0)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/417,495
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/203,866
FILING DATE:
APPLICATION NUMBER: US/07/847,566
FILING DATE:
APPLICATION NUMBER: 07/665,961
FILING DATE: March 7, 1991
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Clark, Paul T.
REGISTRATION NUMBER: 30,162
REFERENCE/DOCKET NUMBER: 00786/119002
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 542-5070
TELEFAX: (617) 542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1389 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)

```

US-08-417-495-2

Query Match 43.2%; Score 622.8; DB 2; Length 1389;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGCACTTGTGTGTGTGCACTGGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGCACTTGTGTGTGTGCACTGGCGCTCTCCCA 60
QY 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 121 TGTACACTTCCCGAGAGAGAGATCAATTCATCGAGAGAGAGAGAGAGAGAGAG 180
DB 121 TGTACACTTCCCGAGAGAGAGATCAATTCATCGAGAGAGAGAGAGAGAGAGAG 180
QY 181 ATTCTGGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATTCAGCTGAATGATCGGCT 240
DB 181 ATTCTGGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATTCAGCTGAATGATCGGCT 240
QY 241 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 361 CTAGTGTGGATTTAGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCACTGTG 420
DB 361 CTAGTGTGGATTTAGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCACTGTG 420
QY 421 CTGACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 CTGACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 AAAAACAATACAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 AAAAACAATACAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 ACCTGACATGCACTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 ACCTGACATGCACTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
QY 601 GTGCTAGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626
DB 601 GTGCTAGCTTTCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 626

```

RESULT 13

US-08-284-391B-2

Sequence 2, Application US/08284391B

Patent No. 5851828

GENERAL INFORMATION:

APPLICANT: Seed, Brian

APPLICANT: Banpour, Babak

APPLICANT: Romeo, Charles

APPLICANT: Kolanus, Waldemar

TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

NUMBER OF SEQUENCES: 53

CORRESPONDENCE ADDRESS:

ADDRESS: 176 Federal Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02110

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS  
 SOFTWARE: FASTSEQ for Windows Version 2.0  
 CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391B

FILING DATE: 02-AUG-1994

CLASSIFICATION: 514

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/195,395

FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566

FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961

FILING DATE: 07-MAR-1991

ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.

REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-428-0200

TELEFAX: 617-428-7045

TELEX:

INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:

LENGTH: 1389 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

MOLECULE TYPE: cDNA

US-08-284-391B-2

Query Match 43.2%; Score 622.8; DB 2; Length 1389;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGCACTTGTGTGTGTGCACTGGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGCACTTGTGTGTGTGCACTGGCGCTCTCCCA 60
QY 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
DB 61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
QY 121 TGTACACTTCCCGAGAGAGAGATCAATTCATCGAGAGAGAGAGAGAGAGAGAG 180
DB 121 TGTACACTTCCCGAGAGAGAGATCAATTCATCGAGAGAGAGAGAGAGAGAGAG 180
QY 181 ATTCTGGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATTCAGCTGAATGATCGGCT 240
DB 181 ATTCTGGGAAATCAGGGCTCTCTTTTAACTAAAGGTCATTCAGCTGAATGATCGGCT 240
QY 241 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
DB 241 GACTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 360
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGTGTGTGTGTGTGTGTGTGTGTGT 360
QY 361 CTAGTGTGGATTTAGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCACTGTG 420
DB 361 CTAGTGTGGATTTAGCTGTGCACTGTGCACTGTGCACTGTGCACTGTGCACTGTG 420
QY 421 CTGACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
DB 421 CTGACCTTTGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 480
QY 481 AAAAACAATACAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
DB 481 AAAAACAATACAGGGGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 540
QY 541 ACCTGACATGCACTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600
DB 541 ACCTGACATGCACTGTCTTTCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 600

```

Qy	601	GTCCTAGCTTCCAGAAGGCCTCCGG	626
Db	601	GTCCTAGCTTCCAGAAGGCTCCAG	626

RESULT 14

```

US-09-218-950-2
: Sequence 2, Application US/09218950
: Patent No. 6284240
:
: GENERAL INFORMATION:
: APPLICANT: Seed, Brian
: APPLICANT: Banapour, Babak
: APPLICANT: Romeo, Charles
: APPLICANT: Kolanus, Waldemar
: TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
: TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS
: NUMBER OF SEQUENCES: 53
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: Clark & Biring LLP
: STREET: 176 Federal Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02110
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM Compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/218,950
: FILING DATE:
: CLASSIFICATION:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US/08/284,391
: FILING DATE: 02-AUG-1994
: APPLICATION NUMBER: 08/195,395
: FILING DATE: 14-FEB-1994
: APPLICATION NUMBER: 07/847,566
: FILING DATE: 06-MAR-1992
: APPLICATION NUMBER: 07/665,961
: FILING DATE: 07-MAR-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Biring, Karen L
: REGISTRATION NUMBER: 35,228
: REFERENCE/DOCKET NUMBER: 00786/247001
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-428-0200
: TELEFAX: 617-428-7045
: TELEX:
: INFORMATION FOR SEQ ID NO: 2:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 1389 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: CDNA
US-09-218-950-2

```

Query Match	43.2%;	Score 622.8;	DB 3;	length 1389;
Best Local Similarity	99.7%;	Pred. No. 1.3e-137;		
Matches 624;	Conservative	0;	Mismatches 2;	Indels 0;
			Gaps	0

Qy	1	ATGAACCGGGAGTCCCTTTTAAAGCACTTGCTTGGTGCTGCAACATGCGCCTCTCCCA	60
Db	1	ATGAACCGGGAGTCCCTTTTAAAGCACTTGCTTGGTGCTGCAACATGCGCCTCTCCCA	60
Qy	61	GCAGCCACTTCAGGGAAGAAGATGTGGCTGGGCAAAAAAGGGATACAGTGGAACTBACC	120
Db	61	GCAGCCACTTCAGGGAAGAAGATGTGGCTGGGCAAAAAAGGGATACAGTGGAACTBACC	120
Qy	121	TGTACAGTCTCCCGAAGAAGACATACATTCACCTGTAATAAACTCCACACGATTAAG	180
Db	121	TGTACAGTCTCCCGAAGAAGACATACATTCACCTGTAATAAACTCCACACGATTAAG	180

Db	121	TGTAAGCTTCCAGAAAGAGCATACATTCCATGGAAAACTCCAACTCAGATTAAG	180
Oy	181	ATTCTGGAAATCAGGGCTCTTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT	240
Db	181	ATTCTGGAAATCAGGGCTCTTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT	240
Oy	241	GACTCAAGAAAGAGCCTTTGGGACCAAGAAACTTCCCCGATCATCAAGAACTTAAAG	300
Db	241	GACTCAAGAAAGAGCCTTTGGGACCAAGAAACTTCCCCGATCATCAAGAACTTAAAG	300
Oy	301	ATAGAAGCTCAGATACATTACATCTGTAGTGAAGAGCAAGAAAGAGGTGCAATTG	360
Db	301	ATAGAAGCTCAGATACATTACATCTGTAGTGAAGAGCAAGAAAGAGGTGCAATTG	360
Oy	361	CTAGTGTGGGATTGACTGCCAATCTGTACACCCACTGCTTCAAGGGCAGAGCTTGACC	420
Db	361	CTAGTGTGGGATTGACTGCCAATCTGTACACCCACTGCTTCAAGGGCAGAGCTTGACC	420
Oy	421	CTAGCCTTGGAGAGCCCCCTGGTGTAGTAAAGCCCTCAGTGCAATGTAGAGATCCAAAGGGGT	480
Db	421	CTAGCCTTGGAGAGCCCCCTGGTGTAGTAAAGCCCTCAGTGCAATGTAGAGATCCAAAGGGGT	480
Oy	481	AAAAACATACAGGGGGGGAAGACCCCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC	540
Db	481	AAAAACATACAGGGGGGGAAGACCCCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC	540
Oy	541	ACCTGACATGCATGTCTTTCAGAAACAGAAAGAGGTGAGTTCAAAATAGACATGTG	600
Db	541	ACCTGACATGCATGTCTTTCAGAAACAGAAAGAGGTGAGTTCAAAATAGACATGTG	600
Oy	601	GTGCTAGCTTTCAGAAAGGCTCCAG	626
Db	601	GTGCTAGCTTTCAGAAAGGCTCCAG	626

**RESULT 15**

```

US-08-394-388A-2
; Sequence 2, Application US/08394388A
; Patent No. 6753162
;
; GENERAL INFORMATION:
; APPLICANT: Seed, Brian
; APPLICANT: Banapour, Babak
; APPLICANT: Romeo, Charles
; APPLICANT: Kolanus, Waldemar
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSD for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/394,388A
; FILING DATE: 24-FEB-1995
; CLASSIFICATION: 514
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
;

```

## ATTORNEY/AGENT INFORMATION:

NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:

## INFORMATION FOR SEQ ID NO: 2:

SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-2

Query Match 43.2%; Score 622.8; DB 4; Length 1389;

Best Local Similarity 99.7%; Pred. No. 1.3e-137;

Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACGGGGAGTCCCTTTTAGCACTTGCTGCTGCAACTGGCGCTCTCCA 60
   |||
Db 1 ATGAACGGGGAGTCCCTTTTAGCACTTGCTGCTGCAACTGGCGCTCTCCA 60

QY 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 120
   |||
Db 61 GCAGCCACTCAGGAAAGAAAGTGTGCTGGCAAAAAAGGGGATACAGTGAATGACC 120

QY 121 TGTACAGCTTCCAGAAAGAGCATACAAATTCACCTGAAAAAACTCCAAACAGATAAG 180
   |||
Db 121 TGTACAGCTTCCAGAAAGAGCATACAAATTCACCTGAAAAAACTCCAAACAGATAAG 180

QY 181 ATTCTGGGAATCAGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 240
   |||
Db 181 ATTCTGGGAATCAGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 240

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGATCTTAAG 300
   |||
Db 241 GACTCAAGAAAGACCTTTGGGACCAAGAAACTTCCCGCTGATCATCAAGATCTTAAG 300

QY 301 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360
   |||
Db 301 ATAGAAAGCTCAGATCTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360

QY 361 CTAGTGTTCGATGACTGCACTGCACTGCAACCCACCTGCTTCAGGGGCAAGCCTGACC 420
   |||
Db 361 CTAGTGTTCGATGACTGCACTGCACTGCAACCCACCTGCTTCAGGGGCAAGCCTGACC 420

QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTCAATGTAGGAATCCAAAGGGGT 480
   |||
Db 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTCAATGTAGGAATCCAAAGGGGT 480

QY 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCAGATATGAGC 540
   |||
Db 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCAGATATGAGC 540

QY 541 ACCTGAACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATCGTG 600
   |||
Db 541 ACCTGAACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATATGACATCGTG 600

QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
   |||
Db 601 GTGCTAGCTTCCAGAGGCTCCAG 626
```

Search completed: September 20, 2005, 10:22:34  
Job time : 300 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 10:22:42 ; Search time 2458 Seconds

(without alignments)  
3903.215 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgacacggagagtcctt.....gtcagcagcataactg 1440

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 7389322 seqs, 3331285599 residues

Total number of hits satisfying chosen parameters: 14778644

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US10\_PUBCOMB.seq:\*
- 23: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*
- 24: /cgn2\_6/ptodata/1/pubpna/US11\_PUBCOMB.seq:\*
- 25: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*
- 26: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	43.4	8911	18	US-10-612-192-3
2	624.4	43.4	1377	14	US-10-103-597A-38
3	624.4	43.4	1377	15	US-10-188-444-38
4	624.4	43.4	1742	14	US-10-151-274-7
5	624.4	43.4	1742	18	US-10-641-643-1013
6	622.8	43.2	1304	10	US-09-939-537-28
7	622.8	43.2	1389	10	US-09-939-537-2

8	622.8	43.2	1389	11	US-09-243-008-2	Sequence 2, Appl1
9	622.8	43.2	1599	10	US-09-939-537-3	Sequence 3, Appl1
10	622.8	43.2	1599	11	US-09-243-008-3	Sequence 3, Appl1
11	622.8	43.2	1728	10	US-09-939-537-1	Sequence 1, Appl1
12	622.8	43.2	1728	11	US-09-243-008-1	Sequence 1, Appl1
13	622.8	43.2	1742	11	US-09-891-118A-8	Sequence 8, Appl1
14	622.8	43.2	1745	16	US-10-024-329-1	Sequence 1, Appl1
15	622.8	43.2	3084	15	US-10-207-655-169	Sequence 169, Appl
16	621.2	43.1	1273	11	US-09-891-118A-1	Sequence 1, Appl1
17	621.2	43.1	1416	15	US-10-157-408-2	Sequence 2, Appl1
18	621.2	43.1	1416	15	US-10-097-044A-2	Sequence 2, Appl1
19	621.2	43.1	1416	15	US-10-769-247-2	Sequence 2, Appl1
20	613.4	42.6	1796	8	US-08-485-163-2	Sequence 2, Appl1
21	613.4	42.6	1796	9	US-09-766-995-1	Sequence 1, Appl1
22	612.2	42.5	1149	8	US-08-485-163-6	Sequence 6, Appl1
23	612.2	42.5	1149	9	US-09-766-995-5	Sequence 5, Appl1
24	612.2	42.5	3273	17	US-10-397-565-7	Sequence 7, Appl1
25	612.2	42.5	11228	17	US-10-397-565-6	Sequence 6, Appl1
26	612	42.5	617	20	US-10-493-676-7	Sequence 4, Appl1
27	612	42.5	2482	8	US-08-485-163-4	Sequence 4, Appl1
28	612	42.5	2482	9	US-09-766-995-3	Sequence 3, Appl1
29	610	42.4	1350	20	US-10-493-676-10	Sequence 10, Appl1
30	610	42.4	1353	20	US-10-493-676-5	Sequence 5, Appl1
31	609	42.3	609	19	US-10-466-836-24	Sequence 24, Appl1
32	607.4	42.2	609	22	US-10-513-725-9	Sequence 9, Appl1
33	607.4	42.2	609	22	US-10-512-966-6	Sequence 6, Appl1
34	607.4	42.2	719	10	US-09-939-537-30	Sequence 30, Appl1
35	571.4	39.7	573	16	US-10-024-329-31	Sequence 31, Appl1
36	552.2	38.3	1508	15	US-10-157-408-5	Sequence 5, Appl1
37	552.2	38.3	1508	15	US-10-097-044A-5	Sequence 5, Appl1
38	552.2	38.3	1508	19	US-10-769-247-5	Sequence 5, Appl1
39	549.4	38.2	1113	9	US-09-759-841-5	Sequence 5, Appl1
40	542.4	37.7	563	14	US-10-794-774-24	Sequence 24, Appl1
41	542.4	37.7	563	21	US-10-073-118-25	Sequence 25, Appl1
42	541.8	37.6	2319	14	US-10-073-118-25	Sequence 25, Appl1
43	541.8	37.6	2319	21	US-10-794-774-25	Sequence 25, Appl1
44	535.6	37.2	1759	9	US-09-934-060A-12	Sequence 12, Appl1
45	535.6	37.2	2159	9	US-09-934-060A-1	Sequence 1, Appl1

#### ALIGNMENTS

RESULT 1  
US-10-612-192-3  
Sequence 3, Application US/10612192  
Publication No. US20040076636A1  
GENERAL INFORMATION:  
APPLICANT: Pal, Ranajit  
APPLICANT: Markham, Phillip  
APPLICANT: Keen, Timothy  
APPLICANT: Whitney, Stephen  
APPLICANT: Kalyanaraman, V.S.  
TITLE OF INVENTION: HIV Immunogenic Complexes  
FILE REFERENCE: 00711 CIP  
CURRENT APPLICATION NUMBER: US/10/612,192  
CURRENT FILING DATE: 2003-07-02  
PRIOR APPLICATION NUMBER: US 09/905,962  
PRIOR FILING DATE: 2001-07-17  
PRIOR APPLICATION NUMBER: US 09/479,675  
PRIOR FILING DATE: 2000-01-07  
PRIOR APPLICATION NUMBER: US 09/075,544  
PRIOR FILING DATE: 1998-05-11  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.2  
SEQ ID NO 3  
LENGTH: 8911  
TYPE: DNA  
ORGANISM: Artificial  
FEATURE:  
OTHER INFORMATION: Expression Vector PTK13+Neo4  
US-10-612-192-3



US-10-188-444-38

Query Match 43.4%; Score 624.4; DB 15; Length 1377;

Best Local Similarity 99.8%; Pred. No. 6.6e-11; Indels 0; Gaps 0;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTATGACCTTCTGTGTGTCGAATGCGCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTATGACCTTCTGTGTGTCGAATGCGCTCTCCCA 60
QY 61 GCAGCCACTCAGGAGAAAGTGTGTGTGCGCAAAAAGGGATACATGAACTGACC 120
DB 61 GCAGCCACTCAGGAGAAAGTGTGTGTGCGCAAAAAGGGATACATGAACTGACC 120
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAACCTCAACCGATTAAG 180
DB 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAACCTCAACCGATTAAG 180
QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240
QY 241 GACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCATTA 360
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCATTA 360
QY 361 CTAGTGTGGATTTGCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
DB 361 CTAGTGTGGATTTGCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGATGATGAGATCCAGGGGT 480
DB 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGATGATGAGATCCAGGGGT 480
QY 481 AAAAACAATCAGGGGGGAGAACCTCTCGTGTCTGACCTGAGCTCCAGATATGTCG 540
DB 481 AAAAACAATCAGGGGGGAGAACCTCTCGTGTCTGACCTGAGCTCCAGATATGTCG 540
QY 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATG 600
DB 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATG 600
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
```

RESULT 4

US-10-151-274-7

; Sequence 7, Application US/10151274

; Publication No. US20030064071A1

; GENERAL INFORMATION:

; APPLICANT: Liltman, Dan R.

; APPLICANT: Kwon, Douglas S.

; APPLICANT: van Kooyk, Yvette

; APPLICANT: Gelfendenbeck, Theo

; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY

; TITLE OF INVENTION: INTD

; FILE REFERENCE: 1049-1-017

; CURRENT APPLICATION NUMBER: US/10/151.274

; PRIOR APPLICATION NUMBER: 2002-05-20

; PRIOR FILING DATE: US/09/517,605

; NUMBER OF SEQ ID NOS: 17

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 7

; LENGTH: 1742

; TYPE: DNA

; ORGANISM: Homo sapiens

US-10-151-274-7

Query Match 43.4%; Score 624.4; DB 14; Length 1742;

Best Local Similarity 99.8%; Pred. No. 7.1e-17; Indels 0; Gaps 0;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
QY 1 ATGAACCGGGAGTCCCTTTTATGACCTTCTGTGTGTCGAATGCGCTCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTATGACCTTCTGTGTGTCGAATGCGCTCTCCCA 135
QY 61 GCAGCCACTCAGGAGAAAGTGTGTGTGCGCAAAAAGGGATACATGAACTGACC 120
DB 136 GCAGCCACTCAGGAGAAAGTGTGTGTGCGCAAAAAGGGATACATGAACTGACC 195
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAACCTCAACCGATTAAG 180
DB 196 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAACCTCAACCGATTAAG 255
QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 256 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 315
QY 241 GACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCATTA 360
DB 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCATTA 435
QY 361 CTAGTGTGGATTTGCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
DB 436 CTAGTGTGGATTTGCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 495
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGATGATGAGATCCAGGGGT 555
DB 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGATGATGAGATCCAGGGGT 555
QY 481 AAAAACAATCAGGGGGGAGAACCTCTCGTGTCTGACCTGAGCTCCAGATATGTCG 540
DB 556 AAAAACAATCAGGGGGGAGAACCTCTCGTGTCTGACCTGAGCTCCAGATATGTCG 615
QY 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATG 600
DB 616 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGTGAAGTTCAAAATAGACATG 675
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCGG 701
```

RESULT 5

US-10-641-643-1013

; Sequence 1013, Application US/10641643

; Publication No. US20040077003A1

; GENERAL INFORMATION:

; APPLICANT: Cocks, Benjamin G.

; APPLICANT: Susan G. Stuart

; APPLICANT: Jeffrey J. Sellhammer

; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL

; GENE EXPRESSION

; NUMBER OF SEQUENCES: 1508

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.

; STREET: 3174 PORTER DRIVE

; CITY: PALO ALTO

; STATE: CALIFORNIA

; COUNTRY: USA

; ZIP: 94304

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: <Unknown>  
FILING DATE: <Unknown>  
ATTORNEY/AGENT INFORMATION:  
NAME: Zeller, Karen J.  
REGISTRATION NUMBER: 37,071  
REFERENCE/DOCKET NUMBER: PA-0001 US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (650) 855-0555  
TELEFAX: (650) 845-4166  
INFORMATION FOR SEQ ID NO: 1013:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1742 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
IMMEDIATE SOURCE:  
LIBRARY: GENBANK  
CLONE: g179143  
SEQUENCE DESCRIPTION: SEQ ID NO: 1013 :  
US-10-641-643-1013

Query Match 43.4%; Score 624.4; DB 18; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 7.1e-171;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGCTGCACTGGCGCTCTCCCA 60  
DB 76 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGCTGCACTGGCGCTCTCCCA 135  
QY 61 GCAGCCACTAGAGAAAGAGTGTCTGGCAAAAAGGGATACGTGAACTGACC 120  
DB 136 GCAGCCACTAGAGAAAGAGTGTCTGGCAAAAAGGGATACGTGAACTGACC 195  
QY 121 TGTACGCTTCCAGAAAGAGATCAATTCACACTGGAATACTCCACGATTAAG 180  
DB 196 TGTACGCTTCCAGAAAGAGATCAATTCACACTGGAATACTCCACGATTAAG 255  
QY 181 ATTCTGGAAATCAGGGCTCTTTTAACTAAAGTTCATCAAGTGAATCGCGCT 240  
DB 256 ATTCTGGAAATCAGGGCTCTTTTAACTAAAGTTCATCAAGTGAATCGCGCT 315  
QY 241 GACTCAAGAAAGACCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 316 GACTCAAGAAAGACCTTTGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGCAGAAAGAGAGTGAATTG 360  
DB 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGAGCAGAAAGAGAGTGAATTG 435  
QY 361 CTAGTGTTCGATTTGACTGCAACTGTGACACCACTGCTTCAAGGGGCAAGCTGACC 420  
DB 436 CTAGTGTTCGATTTGACTGCAACTGTGACACCACTGCTTCAAGGGGCAAGCTGACC 495  
QY 421 CTGACCTTTGAGAGCCCTTGGTAGAGCCCTCACTGATGAAATGAGATCCAAAGGGT 480  
DB 496 CTGACCTTTGAGAGCCCTTGGTAGAGCCCTCACTGATGAAATGAGATCCAAAGGGT 555  
QY 481 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGGAGCTCCAGATAGTGC 540  
DB 556 AAAAACAATACAGGGGGAGAGACCTCTCCGTGTCTCAGCTGGAGCTCCAGATAGTGC 615  
QY 541 ACCTGGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATGACATCTGT 600  
DB 616 ACCTGGACATGCACTGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATGACATCTGT 675  
QY 601 GTCTAGCTTTCCAGAAAGGCTCCGG 626

DB 676 GTCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 6  
US-09-939-537-28  
Sequence 28, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 28:  
US-09-939-537-28

Query Match 43.2%; Score 622.8; DB 10; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTAGAGAAAGAGTGTCTGGCAAAAAGGGATACGTGAACTGACC 120  
DB 171 GCAGCCACTAGAGAAAGAGTGTCTGGCAAAAAGGGATACGTGAACTGACC 230  
QY 121 TGTACGCTTCCAGAAAGAGATCAATTCACACTGGAATACTCCACGATTAAG 180

Db 231 TGTACAGCTTCCAGAGAAAGACATCAATTCACCTGAAACCTCCAAACGAGATAAG 290  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTACTAAGAGTCCATCAAGTGAATGCGCT 240  
 Db 291 ATTCTGGGAATCAGGGCTCTTCTTACTAAGAGTCCATCAAGTGAATGCGCT 350  
 QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300  
 Db 351 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 410  
 QY 301 ATGAAAGACTCAATCTTCACTGTGAAAGTGAAGACCAAGAGAGAGTGCATTG 360  
 Db 411 ATGAAAGACTCAATCTTCACTGTGAAAGTGAAGACCAAGAGAGAGTGCATTG 470  
 QY 361 CTAGTGTGGATTTGACTGCAACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420  
 Db 471 CTAGTGTGGATTTGACTGCAACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 530  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
 Db 531 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 590  
 QY 481 AAAAATCAGAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGAGCTCCAGATAGTGC 540  
 Db 591 AAAAATCAGAGGGGGGAGAGACCTCTCTGTGTCTGAGCTGAGCTCCAGATAGTGC 650  
 QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 600  
 Db 651 ACCTGACATGACCTGTCTTGCAGAACCAAGAGAGTGAAGTTCAAAATAGACATGTC 710  
 QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626  
 Db 711 GTGCTAGCTTTCCAGAGGCTCCAG 736

RESULT 7  
 US-09-939-537-2  
 ; Sequence 2, Application US/09939537  
 ; Publication No. US20030138410A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Seed, Brian  
 ; Banapour, Babak  
 ; Romeo, Charles  
 ; Kolanus, Waldemar  
 ; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
 ; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
 ; NUMBER OF SEQUENCES: 53  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESS: Clark & Elbing LLP  
 ; STREET: 176 Federal Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/939,537  
 ; FILING DATE: 24-Aug-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/284,391  
 ; FILING DATE: 02-AUG-1994  
 ; APPLICATION NUMBER: 08/195,395  
 ; FILING DATE: 14-FEB-1994  
 ; APPLICATION NUMBER: 07/847,566  
 ; FILING DATE: 06-MAR-1992  
 ; APPLICATION NUMBER: 07/665,961  
 ; FILING DATE: 07-MAR-1991  
 ; ATTORNEY/AGENT INFORMATION:

; NAME: Elbing, Karen L  
 ; REGISTRATION NUMBER: 35,238  
 ; REFERENCE/DOCKET NUMBER: 00786/247001  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-428-0200  
 ; TELEFAX: 617-428-7045  
 ; TELEX: <Unknown>  
 ; INFORMATION FOR SEQ ID NO: 2:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 1389 base pairs  
 ; TYPE: nucleic acid  
 ; STRANDEDNESS: double  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: cDNA  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
 US-09-939-537-2

Query Match 43.2%; Score 622.8; DB 10; Length 1389;  
 Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGGAGTCCCTTTTAGCACTGCTCTGTGCTGCACTGGGCGCTCCCA 60  
 Db 1 ATGAACCGGGGAGTCCCTTTTAGCACTGCTCTGTGCTGCACTGGGCGCTCCCA 60  
 QY 61 GCAGCCACTCAGGGGAAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 Db 61 GCAGCCACTCAGGGGAAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 QY 121 TGTACAGCTTCCAGAGAGAGCATATTCCTGAGAAAATCTCAACAGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAGAGCATATTCCTGAGAAAATCTCAACAGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTACTAAGAGTCCATCCAGCTGAATGATCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTCTTACTAAGAGTCCATCCAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300  
 Db 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAG 300  
 QY 301 ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
 Db 301 ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
 QY 361 CTAGTGTGGATTTGACTGCAACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420  
 Db 361 CTAGTGTGGATTTGACTGCAACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
 QY 481 AAAAATCAGAGGGGGGAGAGACCTCTCTGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
 Db 481 AAAAATCAGAGGGGGGAGAGACCTCTCTGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
 QY 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGTGAAGTCAAAATAGACATGTC 600  
 Db 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGTGAAGTCAAAATAGACATGTC 600  
 QY 601 GTGCTAGCTTTCCAGAGGCTCCGG 626  
 Db 601 GTGCTAGCTTTCCAGAGGCTCCAG 626

RESULT 8  
 US-09-243-008-2  
 ; Sequence 2, Application US/09243008  
 ; Publication No. US20040005334A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Seed, Brian et al.  
 ; TITLE OF INVENTION: Redirection of Cellular Immunity by

Receptor Chimeras  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
OPERATING SYSTEM: IBM PC/2 Model 502 or 55SX  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,176  
FILING DATE: SEPTEMBER 11, 1995  
APPLICATION NUMBER: 08/203,866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 6, 1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-243-008-2

Query Match 43.2%; Score 622.8; DB 11; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.9e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
1 ATGAACCGGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60  
61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120  
61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTACC 120  
121 TGTAAGCTTCCGAGAAAGAGCATTAATTCACCTGGAATACTCCACAGATTAAG 180  
121 TGTAAGCTTCCGAGAAAGAGCATTAATTCACCTGGAATACTCCACAGATTAAG 180  
181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTCCATCCAAAGCTGAATGTCGGCT 240  
181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGGTCCATCCAAAGCTGAATGTCGGCT 240  
241 GACTCAAGAAAGACCTTTGGGACCAAGAACTCCCGCTGATCATCAAGATCTTAAG 300  
241 GACTCAAGAAAGACCTTTGGGACCAAGAACTCCCGCTGATCATCAAGATCTTAAG 300  
301 ATGAAGACTCAGATCTTAATCTGTGAAGTGAAGGACCAAGAAAGAGTGAATTG 360  
301 ATGAAGACTCAGATCTTAATCTGTGAAGTGAAGGACCAAGAAAGAGTGAATTG 360  
361 CTAGTGTGGGATTTGACTGCCAACTGACACCCCACTGCTTCAAGGGGCGAGGCTGACC 420

361 CTAGTGTGGGATTTGACTGCCAACTGACACCCCACTGCTTCAAGGGGCGAGGCTGACC 420  
421 CTGACCTTGAAGACCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCAAGAGGT 480  
421 CTGACCTTGAAGACCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCAAGAGGT 480  
481 AAAACATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
481 AAAACATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
541 ACCTGACATGACCTCTTTCAGAGAACCAAGAGTGAAGTCAATTAAGCATCTG 600  
541 ACCTGACATGACCTCTTTCAGAGAACCAAGAGTGAAGTCAATTAAGCATCTG 600  
601 GTGCTAGCTTTCAGAGAGCTCCGG 626  
601 GTGCTAGCTTTCAGAGAGCTCCGG 626  
601 GTGCTAGCTTTCAGAGAGCTCCGG 626

RESULT 9  
US-09-939-537-3  
Sequence 3, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1599 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:

US-09-939-537-3

Query Match 43.2%; Score 622.8; DB 10; Length 1599;  
 Best Local Similarity 99.7%; Pred. No. 2e-170; Indels 0; Gaps 0;  
 Matches 624; Conservative 0; Mismatches 2;

QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 Db 61 GCAGCCACTCAGGGAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCAGAGAGAGATCAATTCACGGAATACTCCACGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAGAGATCAATTCACGGAATACTCCACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAGAAAGCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAGAAAGCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACGAGAGAGAGTGCATTAAG 360  
 Db 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACGAGAGAGAGTGCATTAAG 360  
 QY 361 CTAGTGTCCGATTTGACTGCACTGTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 420  
 Db 361 CTAGTGTCCGATTTGACTGCACTGTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAAGTCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAAGTCCAAAGGGGT 480  
 QY 481 AAAAACAATAAGGGGGGAGAGACCTCTCGTGTCTGAGTGCATCCAGATATGAGGC 540  
 Db 481 AAAAACAATAAGGGGGGAGAGACCTCTCGTGTCTGAGTGCATCCAGATATGAGGC 540  
 QY 541 ACCTGACATGCACTGTCTTGGAGAACGAGAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 Db 541 ACCTGACATGCACTGTCTTGGAGAACGAGAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 QY 601 GTGCTAGCTTTCCAGAGAGGCTCCGG 626  
 Db 601 GTGCTAGCTTTCCAGAGAGGCTCCAG 626

RESULT 10  
 US-09-243-008-3  
 Sequence 3, Application US/09243008  
 Publication No. US2004005334A1

GENERAL INFORMATION:  
 APPLICANT: Seed, Brian et al.

TITLE OF INVENTION: Redirection of Cellular Immunity by  
 Receptor Chimeras

NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston

STATE: MA  
 COUNTRY: USA

ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

COMPUTER: IBM PS/2 Model 502 or 55SX  
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 SOFTWARE: Wordperfect (Version 5.0)

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/243,008  
 FILING DATE: 02-Feb-1999  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/394,176  
 FILING DATE: SEPTEMBER 11, 1995  
 APPLICATION NUMBER: 08/203,866  
 FILING DATE: February 28, 1994  
 APPLICATION NUMBER: 07/847,566  
 FILING DATE: March 6, 1992  
 APPLICATION NUMBER: 07/665,961  
 FILING DATE: March 7, 1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Karen F. Lech, Ph.D.  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/270001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 INFORMATION FOR SEQ ID NO: 3:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1599 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
 US-09-243-008-3

Query Match 43.2%; Score 622.8; DB 11; Length 1599;  
 Best Local Similarity 99.7%; Pred. No. 2e-170; Indels 0; Gaps 0;  
 Matches 624; Conservative 0; Mismatches 2;

QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTCTGCTGCTGCAACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 Db 61 GCAGCCACTCAGGGAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCAGAGAGAGATCAATTCACGGAATACTCCACGATTAAG 180  
 Db 121 TGTACAGCTTCCAGAGAGAGATCAATTCACGGAATACTCCACGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAGAAAGCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAGAAAGCTTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACGAGAGAGAGTGCATTAAG 360  
 Db 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACGAGAGAGAGTGCATTAAG 360  
 QY 361 CTAGTGTCCGATTTGACTGCACTGTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 420  
 Db 361 CTAGTGTCCGATTTGACTGCACTGTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 420  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAAGTCCAAAGGGGT 480  
 Db 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCACTGATGATGAAGTCCAAAGGGGT 480  
 QY 481 AAAAACAATAAGGGGGGAGAGACCTCTCGTGTCTGAGTGAAGTCCAGATTAAGGC 540  
 Db 481 AAAAACAATAAGGGGGGAGAGACCTCTCGTGTCTGAGTGAAGTCCAGATTAAGGC 540  
 QY 541 ACCTGACATGCACTGTCTTGGAGAACGAGAGAGAGTGAAGTCAAAATAGACATGCTG 600  
 Db 541 ACCTGACATGCACTGTCTTGGAGAACGAGAGAGAGTGAAGTCAAAATAGACATGCTG 600

QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

## RESULT 11

US-09-939-537-1  
Sequence 1, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanue, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1728 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-939-537-1

Query Match 43.2%; Score 622.8; DB 10; Length 1728;  
Best Local Similarity 99.7%; Pred. No. 2.1e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGACCGGGAGGCTTTAGGCACTTGTGCTGCGCAACTGGGCTCCCTCCCA 60  
DB 1 ATGACCGGGAGGCTTTAGGCACTTGTGCTGCGCAACTGGGCTCCCTCCCA 60  
QY GCAGCCACTCAGGAAAGAAAGTGTGCTGCGCAAAAAAGGGGATACAGTGAAGTGAAC 120  
DB GCAGCCACTCAGGAAAGAAAGTGTGCTGCGCAAAAAAGGGGATACAGTGAAGTGAAC 120  
QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGGAATACTCAACAGATTAAG 180

DB 121 TGTACAGCTTCCAGAAAGAGCATCAATTCCACTGGAATACTCAACAGATTAAG 180  
QY 181 ATTCTGGAAATCAGGGCTCTTCTTAACTAAAGGTCAATCCAGAGTGAATGCGGCT 240  
DB 181 ATTCTGGAAATCAGGGCTCTTCTTAACTAAAGGTCAATCCAGAGTGAATGCGGCT 240  
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATGAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGAGGATGCAATTG 360  
DB 301 ATGAGACTCAGATCTTACATCTGTGAAGTGAAGGACCAAGAGAGGATGCAATTG 360  
QY 361 CTAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 420  
DB 361 CTAGTGTGGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 420  
QY 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAGAGGCT 480  
DB 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAGAGGCT 480  
QY 481 AAAAATACAGGGGGGAAAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
DB 481 AAAAATACAGGGGGGAAAGCCCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 540  
QY 541 ACCTGACATGCACTGCTTTCGAGAAACAGAAAGTGAATTCATAATACATCTG 600  
DB 541 ACCTGACATGCACTGCTTTCGAGAAACAGAAAGTGAATTCATAATACATCTG 600  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

## RESULT 12

US-09-243-008-1  
Sequence 1, Application US/09243008  
Publication No. US20040005334A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson P.C.  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 50Z or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: WordPerfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243,008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,176  
FILING DATE: SEPTEMBER 11,1995  
APPLICATION NUMBER: 08/203,866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: March 7, 1991  
APPLICATION NUMBER: 07/665,961  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Iech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1728 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:

US-09-243-008-1

Query Match 43.2%; Score 622.8; DB 11; Length 1728;  
 Best Local Similarity 99.7%; Pred. No. 2,1e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAAGTCCCTTTTAAAGCACTTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 1 ATGAACCGGGAAGTCCCTTTTAAAGCACTTGTGCTGCACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCCTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
 DB 61 GCAGCCCTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
 QY 121 TGTACAGCTTCCGAGAAGAGCATCAATTCACATGGAATACTCAACAGATAAG 180  
 DB 121 TGTACAGCTTCCGAGAAGAGCATCAATTCACATGGAATACTCAACAGATAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 DB 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAAGACTCAGATCTTACATCTGGAAGTGAAGCAAGAGAGAGGTGCAATTG 360  
 DB 301 ATGAAGACTCAGATCTTACATCTGGAAGTGAAGCAAGAGAGAGGTGCAATTG 360  
 QY 361 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAAGGGGCAAGCTGACC 420  
 DB 361 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAAGGGGCAAGCTGACC 420  
 QY 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
 DB 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
 QY 481 AAAAACAATAGAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
 DB 481 AAAAACAATAGAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
 QY 541 ACCTGACATGCACTGTCTTGCAAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600  
 DB 541 ACCTGACATGCACTGTCTTGCAAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

RESULT 13  
 US-09-891-119A-8  
 ; Sequence 8, Application US/09891119A  
 ; Publication No. US20040013683A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Maddon, Paul J.  
 ; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
 ; FILE REFERENCE: 24577-CY-B  
 ; CURRENT APPLICATION NUMBER: US/09/891,119A  
 ; CURRENT FILING DATE: 2001-06-25

NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 8  
 ; LENGTH: 1742  
 ; TYPE: DNA  
 ; ORGANISM: Human  
 US-09-891-119A-8

Query Match 43.2%; Score 622.8; DB 11; Length 1742;  
 Best Local Similarity 99.7%; Pred. No. 2,1e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAAGTCCCTTTTAAAGCACTTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAAGTCCCTTTTAAAGCACTTGTGCTGCACTGGCGCTCTCCCA 135  
 QY 61 GCAGCCCTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
 DB 136 GCAGCCCTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 195  
 QY 121 TGTACAGCTTCCGAGAAGAGCATCAATTCACATGGAATACTCAACAGATAAG 180  
 DB 196 TGTACAGCTTCCGAGAAGAGCATCAATTCACATGGAATACTCAACAGATAAG 255  
 QY 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 240  
 DB 256 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGTGAATGATCGGCT 315  
 QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 DB 316 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
 QY 301 ATGAAGACTCAGATCTTACATCTGGAAGTGAAGCAAGAGAGAGGTGCAATTG 360  
 DB 376 ATGAAGACTCAGATCTTACATCTGGAAGTGAAGCAAGAGAGAGGTGCAATTG 435  
 QY 361 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAAGGGGCAAGCTGACC 420  
 DB 436 CTAGTGTGGATTTGACTGCACTGCAACCCAGCTGCTTCAAGGGGCAAGCTGACC 495  
 QY 421 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 480  
 DB 496 CTGACCTTGAAGAGCCCTCTGTGATGAGCCCTCACTGATGATGAGATCCAGGGGT 555  
 QY 481 AAAAACAATAGAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
 DB 556 AAAAACAATAGAGGGGGGAGAACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 615  
 QY 541 ACCTGACATGCACTGTCTTGCAAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 600  
 DB 616 ACCTGACATGCACTGTCTTGCAAGAACAGAGAGGTGAGTTCAAAATAGACATCGTG 675  
 QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 14  
 US-10-024-329-1  
 ; Sequence 1, Application US/10024329  
 ; Publication No. US20030157063A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SANHADJI, Kamel  
 ; APPLICANT: TOURAINE, Jean-Louis  
 ; APPLICANT: LEROY, Pierre  
 ; APPLICANT: MENTALI, Maïd  
 ; TITLE OF INVENTION: Gene therapy using anti-gp41 antibody and cd4 immunoadhesin  
 ; FILE REFERENCE: 109993  
 ; CURRENT APPLICATION NUMBER: US/10/024,329  
 ; CURRENT FILING DATE: 2001-12-21  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1

LENGTH: 1745  
TYPE: DNA  
ORGANISM: human SCD4  
US-10-024-329-1

Query Match 43.2%; Score 622.8; DB 16; Length 1745;  
Best Local Similarity 99.7%; Pred. No. 2.1e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60
Db ATGAACGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 138
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACATGGAATGACC 120
Db GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACATGGAATGACC 198
QY 121 TGTACGCTTCCAGAAAGAGCATACATTCACAGTGAATACTCAACGATTAAG 180
Db TGTACGCTTCCAGAAAGAGCATACATTCACAGTGAATACTCAACGATTAAG 258
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 240
Db ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 318
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATTTAAG 300
Db GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATTTAAG 378
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 360
Db ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 438
QY 361 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420
Db CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 498
QY 421 CTGACCTTGAAGACCCCTGGTGAAGTACCCCTCAGTGCATGTGGAATCCAAGGGGT 480
Db CTGACCTTGAAGACCCCTGGTGAAGTACCCCTCAGTGCATGTGGAATCCAAGGGGT 558
QY 481 AAAAACAATACAGGGGGGAGAACCCCTGCGTGTCTAGCTGGAGCTCCAGATAGTGGC 540
Db AAAAACAATACAGGGGGGAGAACCCCTGCGTGTCTAGCTGGAGCTCCAGATAGTGGC 618
QY 541 ACCTGACATGCACTGTCTTGCAAAACAGAAAGGTGAGTTCAAAATAGACATGCTG 600
Db ACCTGACATGCACTGTCTTGCAAAACAGAAAGGTGAGTTCAAAATAGACATGCTG 678
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db GTGCTAGCTTTCCAGAAAGGCTCCAG 704
```

RESULT 15  
US-10-207-655-169  
Sequence 169, Application US/10207655  
Publication No. US20030118592A1  
GENERAL INFORMATION:  
APPLICANT: Ledbetter, Jeffrey A.  
APPLICANT: Hayden-Ledbetter, Martha S.  
TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
FILE REFERENCE: 390069.401C1  
CURRENT APPLICATION NUMBER: US/10/207,655  
CURRENT FILING DATE: 2002-07-25  
NUMBER OF SEQ. ID NOS: 426  
SOFTWARE: PatentIn version 3.0  
SEQ. ID NO 169  
LENGTH: 3084  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-10-207-655-169

Query Match 43.2%; Score 622.8; DB 15; Length 3084;  
Best Local Similarity 99.7%; Pred. No. 2.5e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 ATGAACGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 60
Db ATGAACGGGAGTCCCTTTTAAAGCACTTCTGTGCTGCAACTGGCGCTCTCCCA 212
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACATGGAATGACC 120
Db GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACATGGAATGACC 272
QY 121 TGTACGCTTCCAGAAAGAGCATACATTCACAGTGAATACTCAACGATTAAG 180
Db TGTACGCTTCCAGAAAGAGCATACATTCACAGTGAATACTCAACGATTAAG 332
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 240
Db ATTCTGGGAATCAGGGCTCTTCTTAATAAGTTCATCAAGCTGAATGATCGGCT 392
QY 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATTTAAG 300
Db GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATTTAAG 452
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 360
Db ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGGAGGTGCAATTG 512
QY 361 CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 420
Db CTAGTGTGGATTTGACTGCAACTCTGACACCACTGCTTCAAGGGGAGAGCTGACC 572
QY 421 CTGACCTTGAAGACCCCTGGTGAAGTACCCCTCAGTGCATGTGGAATCCAAGGGGT 480
Db CTGACCTTGAAGACCCCTGGTGAAGTACCCCTCAGTGCATGTGGAATCCAAGGGGT 632
QY 481 AAAAACAATACAGGGGGGAGAACCCCTGCGTGTCTAGCTGGAGCTCCAGATAGTGGC 540
Db AAAAACAATACAGGGGGGAGAACCCCTGCGTGTCTAGCTGGAGCTCCAGATAGTGGC 692
QY 541 ACCTGACATGCACTGTCTTGCAAAACAGAAAGGTGAGTTCAAAATAGACATGCTG 600
Db ACCTGACATGCACTGTCTTGCAAAACAGAAAGGTGAGTTCAAAATAGACATGCTG 752
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db GTGCTAGCTTTCCAGAAAGGCTCCAG 778
```

Search completed: September 20, 2005, 16:05:03  
Job time : 2461 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 20, 2005, 08:29:42 ; Search time 5142 Seconds  
(without alignments)  
10659.772 Million cell updates/sec

Title: US-09-936-702-4  
Perfect score: 1440  
Sequence: 1 atgaaccggggagtccttc.....gtcagcagctacataactcg 1440

Scoring table: IDENTITY NUC  
Gap 10.0 , Gape 1.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
EST:\*  
1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_gss1:\*  
9: gb\_gss2:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	622.8	43.2	783	6	CD609192 56069931J
2	622.8	43.2	932	5	BX438745 BX438745
3	610.2	42.4	816	4	BI838409 603083373
4	599.4	41.6	822	5	BX457159 BX457159
5	598.8	41.6	801	4	BI822118 603039813
6	589	40.9	840	4	BI838893 603087279
7	577.4	40.1	791	4	BI819727 603041415
8	555.4	38.6	557	7	CV029573 8403_Full
9	539	37.4	713	5	BX437619 BX437619
10	519.4	36.1	723	4	BI915265 603184313
11	506.4	35.2	679	6	CD609191 56069931H
12	504	35.0	732	1	AU141298 AU141298
13	488.8	33.9	754	7	CO246446 AGENCOURT
14	488.4	33.9	710	7	CF125444 UT-HF-ElO
15	483.8	33.6	694	6	CB052686 NISC_g109
16	480.6	33.4	803	7	CO249540 AGENCOURT
17	466	32.4	809	7	CO24894 AGENCOURT
18	457.4	31.8	748	5	BX378811 BX378811
19	452.8	31.4	821	1	CO245985 AGENCOURT
20	437.8	30.4	584	5	BP303735 BP303735
21	435.8	30.3	580	5	BP345966 BP345966
22	429.4	29.8	616	7	CF125140 UT-HF-ElO
23	416.4	28.9	584	5	BP287350 BP287350
24	416.4	28.9	585	5	BP286919 BP286919

C 25	407.6	28.3	751	7	CO248544 AGENCOURT
26	402.6	28.0	687	2	AW752367 RC2-CT020
27	386.8	26.9	583	5	BP361417 BP361417
28	386.4	26.8	582	5	BP339950 BP339950
29	385.8	26.8	583	5	BP287803 BP287803
30	383.4	26.6	579	5	BP345874 BP345874
31	383.4	26.6	582	5	BP300162 BP300162
32	374.6	26.0	767	1	AU136822 AU136822
33	366.6	25.5	686	6	CB052685 NISC_g109
34	364.2	25.3	755	7	CO246525 AGENCOURT
35	358.4	24.9	564	5	BP277052 BP277052
36	354.8	24.6	638	2	AW752358 RC2-CT020
37	345	24.0	893	1	AL553628 AL553628
38	328.8	22.8	707	7	CK835626 4060108 B
39	325.4	22.6	581	5	BP367248 BP367248
40	324.8	22.6	582	5	BP344125 BP344125
41	302.8	21.0	567	6	CB997693 AGENCOURT
42	299.4	20.8	583	5	BP344191 BP344191
43	261.2	18.1	336	2	AW407282 UT-HF-BLO
44	261.2	18.1	374	2	AW406380 UT-HF-BLO
45	261.2	18.1	733	4	BG685053 BG685053

## ALIGNMENTS

RESULT 1  
LOCUS CD609192 783 bp mRNA linear EST 12-JAN-2004  
DEFINITION 56069931J1 FLP Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD609192  
VERSION CD609192.1 GI:40257455  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.  
1 (bases 1 to 783)  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
TITLE Circular rapid amplification of cDNA ends for high-throughput  
extension cloning of partial genes  
JOURNAL Genomics 84 (1), 205-210 (2004)  
COMMENT Contact: Fu GK  
Incyte Genomics, Inc.  
3160 Porter Dr., Palo Alto, CA 94304, USA  
Tel: 6508454102  
Email: gfu@incyte.com.

FEATURES  
Source Location/Qualifiers  
1..783  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="FLP"  
/note="Vector: pDrive Cloning Vector"

## ORIGIN

Query Match 43.2%; Score 622.8; DB 6; Length 783;  
Best Local Similarity 99.7%; Pred. No. 2.5e-152;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy	1	ATGACCGGAGATCCCTTTAGGCACTTCTGTGCTGTGCACTGGCGCTCTCCCA	60
Db	726	ATGAACCGGAGATCCCTTTAGGCACTTCTGTGCTGTGCACTGGCGCTCTCCCA	667
Qy	61	GCAAGCACTCAGGAAAGAAAGTGCTGCGGCAAAAAGGGATACAGTGAATGACC	120
Db	666	GCAAGCACTCAGGAAAGAAAGTGCTGCGGCAAAAAGGGATACAGTGAATGACC	607
Qy	121	TGACAGCTTCCAGAAAGAGATACATTCATGTGAAAATCCCAACGATGAAG	180
Db	606	TGACAGCTTCCAGAAAGAGATACATTCATGTGAAAATCCCAACGATGAAG	547
Qy	181	ATTCTGGGAATCAGGCTCTTTACTAAAGGTCATCAAGCTGAATGATCGCGCT	240

|||||  
Db 546 ATCTGGGAATCGGGCTCTCTTCTTACTTAAAGGCTCATCCAGCTGAATGATCGGCT 487  
|||  
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300  
|||  
Db 486 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 427  
|||  
Qy 301 ATAGAACTCAGTACTTACTTATCTGTGAAGTGGAGCAAGAAAGAGGTGCATTG 360  
|||  
Db 426 ATAGAACTCAGTACTTACTTATCTGTGAAGTGGAGCAAGAAAGAGGTGCATTG 367  
|||  
Qy 361 CTAGTGTCCGATGTAAGTGCCTGCACTGCAACCCACCTGCTTCCAGGGGAGAGCTTACC 420  
|||  
Db 366 CTAGTGTCCGATGTAAGTGCCTGCACTGCAACCCACCTGCTTCCAGGGGAGAGCTTACC 307  
|||  
Qy 421 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGATCCAAAGGGT 480  
|||  
Db 306 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGCAATGTAGATCCAAAGGGT 247  
|||  
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
|||  
Db 246 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 187  
|||  
Qy 541 ACCTGGACATGCATCTGTCTTGAGACCAAGCAAGAAAGGTGAGTTCAAAATAGATCTG 600  
|||  
Db 186 ACCTGGACATGCATCTGTCTTGAGACCAAGCAAGAAAGGTGAGTTCAAAATAGATCTG 127  
|||  
Qy 601 GTGCTAGCTTTCCAGAAAGCCTCCGG 626  
|||  
Db 126 GTGCTAGCTTTCCAGAAAGCCTCCAG 101  
|||  
RESULT 2  
BX438745 932 bp mRNA linear EST 05-MAY-2004  
LOCUS BX438745 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE002Y001  
DEFINITION 5-PRIME, mRNA sequence.  
ACCESSION BX438745  
VERSION BX438745.2 GI:47035651  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 932)  
AUTHORS Li, W. B., Gruber, C., Jessee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On May 15, 2003 this sequence version replaced gi:30787731.  
Contact: Genoscope  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremlieux, CP 5706 - 91057 Evry cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
was not normalized. Library was constructed by Life Technologies, a  
division of Invitrogen.  
This sequence belongs to sequence cluster 6485.r  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdna7s=CS0DE002AD01Q1&c=6485.r.  
FEATURES  
source  
Location/Qualifiers  
1..932  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0DE002Y001"  
/issue\_type="PLACENTA"  
/tissue\_type="PLACENTA"  
/clone\_id="Homo sapiens PLACENTA"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
with a NotI-oligo(dT) primer. Five prime end enriched,  
double-strand cDNA was digested with Not I and cloned into  
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

ORIGIN Library was not normalized."  
Query Match 43.2%; Score 622.8; DB 5; Length 932;  
Best local Similarity 99.7%; Pred. No. 2,6e-152;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
|||  
Qy 1 ATGACCCGGGAGTCCCTTTTAGCACTTGTCTGTCTGCACTGCAGCTCTCCCA 60  
|||  
Db 180 ATGACCCGGGAGTCCCTTTTAGCACTTGTCTGTCTGCACTGCAGCTCTCCCA 239  
|||  
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTCTGGGCAAAAAGGGGATCACTGAACTGACC 120  
|||  
Db 240 GCAGCCACTCAGGGAAGAAAGTGTCTGGGCAAAAAGGGGATCACTGAACTGACC 299  
|||  
Qy 121 TGTACAGCTTCCAGAAAGACATACATTTCACTGAAAAAACTCCAACAGATAAG 180  
|||  
Db 300 TGTACAGCTTCCAGAAAGACATACATTTCACTGAAAAAACTCCAACAGATAAG 359  
|||  
Qy 181 ATTCTGGAAATCAGGGCTCTTTTAACTAAAGTCTCAAGCTGAATATCGCGCT 240  
|||  
Db 360 ATTCTGGAAATCAGGGCTCTTTTAACTAAAGTCTCAAGCTGAATATCGCGCT 419  
|||  
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGGAATCCCGCTGATCAATCAAGATCTTAAG 300  
|||  
Db 420 GACTCAAGAAAGACCTTTGGGACCAAGGAATCTTCCCTGATCAATCAAGATCTTAAG 479  
|||  
Qy 301 ATAGAACTCAGATACTTACTTACTGTGTGAAGTGAAGACAGAAAGAGAGGTGCATTG 360  
|||  
Db 480 ATAGAACTCAGATACTTACTTACTGTGTGAAGTGAAGACAGAAAGAGAGGTGCATTG 539  
|||  
Qy 361 CTAGTGTCCGATGTAAGTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 420  
|||  
Db 540 CTAGTGTCCGATGTAAGTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGCCTTGC 599  
|||  
Qy 421 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGAGTCAATGTAGATCCAAAGGGT 480  
|||  
Db 600 CTGACCTTGGAGAGCCCGCTGTAGTAGCCCTCACTGAGTCAATGTAGATCCAAAGGGT 659  
|||  
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540  
|||  
Db 660 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 719  
|||  
Qy 541 ACTGGACATGCATCTGTCTTGAGACCAAGCAAGAGGTGAGTTCAAAATAGATCTG 600  
|||  
Db 720 ACTGGACATGCATCTGTCTTGAGACCAAGCAAGAGGTGAGTTCAAAATAGATCTG 779  
|||  
Qy 601 GTGCTAGCTTTCCAGAAAGCCTCCGG 626  
|||  
Db 780 GTGCTAGCTTTCCAGAAAGCCTCCAG 805  
|||  
RESULT 3  
BI838409 816 bp mRNA linear EST 04-OCT-2001  
LOCUS BI838409  
DEFINITION 603083373F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5222323 5',  
mRNA sequence.  
ACCESSION BI838409  
VERSION BI838409.1 GI:15949959  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 816)  
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgaabs-r@mail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LINL at:  
<http://image.jnl.gov>  
 Plate: LAM11559 row: b column: 20  
 High quality sequence stop: 803.  
 Location/Qualifiers

## FEATURES

source

1. 816  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5222323"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 42.4%; Score 610.2; DB 4; Length 816;

Best Local Similarity 99.4%; Pred. No. 5.2e-149; Mismatches 3; Indels 1; Gaps 1;

Matches 623; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

QY 1 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGTGTCGAACTGGCGCTCTCCCA 60

DB 167 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGTGTCGAACTGGCGCTCTCCCA 226

QY 61 GCAGCCCTCAGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAATCTACC 120

DB 227 GCAGCCCTCAGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAATCTACC 286

QY 121 TGTACAGCTTCCGAGAAAGAGCATACATTCACATGAAAACTCCACAGATAAAG 180

DB 287 TGTACAGCTTCCGAGAAAGAGCATACATTCACATGAAAACTCCACAGATAAAG 346

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 240

DB 347 ATTCTGGGAATCAGGGCTCTTCTTAACATAAGGTCATCAAGCTGAATGATCGGCT 406

QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300

DB 407 GACTCAAGAAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 466

QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGGACCAAGAGGAGGTGCAATTG 360

DB 467 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGGACCAAGAGGAGGTGCAATTG 526

QY 361 CTAGTGTTCGATGATCTGCACTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 420

DB 527 CTAGTGTTCGATGATCTGCACTGCAACCCACCTGCTTCAAGGGGAGAGCTGACC 586

QY 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480

DB 587 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 646

QY 481 AAAAACTACAGGGGGGAGAACCTCTCC- GTGTCTCAGCTGAGGCTCCAGATTAATGG 539

DB 647 AAAAACTACAGGGGGGAGAACCTCTCCGATCTCAGCTGAGGCTCCAGATTAATGG 706

QY 540 CACCTGACATGCACTGCTTTCAGAAACAGAGAAGGTGAGTTCAAAATAGACATCGT 599

DB 707 CACCTGACATGCACTGCTTTCAGAAACAGAGAAGGTGAGTTCAAAATAGACATCGT 766

QY 600 GGTGCTAGCTTTCCAGAAAGGCTCTCGG 626

DB 767 GGTGCTAGCTTTCCAGAAAGGCTCTCAG 793

RESULT 4  
 BX457159 822 bp mRNA linear EST 05-MAY-2004  
 LOCUS BX457159  
 DEFINITION BX457159 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YF07  
 5-PRIME, mRNA sequence.  
 ACCESSION BX457159  
 VERSION BX457159.2 GI:47051693  
 KEYWORDS EST.

## SOURCE

ORGANISM

## REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

On May 22, 2003 this sequence version replaced gi:3102083.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web: [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdna?b=CS0CAP006YF07&id=6485.r>.

## FEATURES

source

1. 822  
 Location/Qualifiers

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP006YF07"  
 /issue\_type="THYMUS"  
 /clone\_lib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 41.6%; Score 599.4; DB 5; Length 822;

Best Local Similarity 99.2%; Pred. No. 3.6e-146; Mismatches 2; Indels 2; Gaps 2;

Matches 622; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

QY 1 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGTGTCGAACTGGCGCTCTCCCA 60

DB 183 ATGAACCGGGGAGTCCCTTTAGGCACTTGTGTGTGTCGAACTGGCGCTCTCCCA 242

QY 61 G-CAGCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATCTGAC 119

DB 243 G-CAGCACTCAGGAAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAATCTGAC 302

QY 120 CTGTACAGCTTCCGAGAAAGAGCATACATTCACCTGGAATAAATCCAACTGATTA 179

DB 303 CTGTACAGCTTCCGAGAAAGAGCATACATTCACCTGGAATAAATCCAACTGATTA 362

QY 180 GATTCTGGGAAATCAGGGCTCTTCTTAATTAAGGTCATCCAGCTGAATGATGCGC 239

DB 363 GATTCTGGGAAATCAGGGCTCTTCTTAATTAAGGTCATCCAGCTGAATGATGCGC 422

QY 240 TGACTAAGAAGAGCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGATCTTA 299

DB 423 TGACTAAGAAGAGCTTTGGGACCAAGAAATTCCTCCCTGATCATCAAGATCTTA 482

QY 300 GATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGGTGCAATT 359

DB 483 GATGAAGACTCAGATCTTACATCTGTGAAGTGGAGACCAAGAGAGAGGTGCAATT 542

QY 360 GCTAGTGTGGATTGACTGCCAACTCTGAACCCACCTCTTCCAGGGGCGAGAGCTTGAC 419  
 Db 543 GCTAAGTTCGTGATTGACTCCAACTCTGACACCCACCTCTTCAGGGGCGAGAGCTTGAC 602  
 QY 420 CCTGACCTTGAGAGAGCCCTCTGTTAGTACCCCTGAGTCAATGTAGAGTCCAAAGGGG 479  
 Db 603 CCTGACCTTGAGAGAGCCCTCTGTTAGTACCCCTGAGTCAATGTAGAGTCCAAAGGGG 662  
 QY 480 TAAAAACATACAGAGGGGGGAGAGACCTCTCTGTTCTCACTGAGAGTCCAGAGTATGTGG 539  
 Db 663 TAAAAACATACAGAGGGGGGAGAGACCTCTCTGTTCTCACTGAGAGTCCAGAGTATGTGG 721  
 QY 540 CACTGGAACATGACATCTCTTTCAGAAACCAAGAAAGGTGAGTCAAAATAGACATCGT 599  
 Db 722 CACTGGAACATGACATCTCTTTCAGAAACCAAGAAAGGTGAGTCAAAATAGACATCGT 781  
 QY 600 GGTGCTAGCTTTCCAGAAAGGCTCCGG 626  
 Db 782 GGTGCTAGCTTTCCAGAAAGGCTCCGG 808

RESULT 5  
 BI822118 801 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603039813F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5180642 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI822118  
 VERSION BI822118.1 GI:15933668  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS 1 (bases 1 to 801)  
 JOURNAL NIH-MGC http://mgs.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1450 row: j column: 03  
 High quality sequence stop: 793.

## FEATURES

## Source

1..801  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5180642"  
 /lab\_host="DH10B"  
 /clone\_id="NIH\_MGC\_115"  
 /note="Organ: pooled brain, lung, testis; Vector:  
 pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
 source anonymous pool of 6 male brains, age range 23-27; 1  
 male lung, age 27; and 1 male testis, age 69. Library is  
 oligo-dT primed and directionally cloned (EcoRV site is  
 destroyed upon cloning). Average insert size 1.8 kb,  
 insert size range 1-3 kb. Library is normalized and  
 enriched for full-length clones and was constructed by C.  
 Gruber (Invitrogen). Research Genetics tracking code  
 021. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 41.6%; Score 598.8; DB 4; Length 801;  
 Best Local Similarity 99.4%; Pred. No. 5,1e-146;  
 Matches 622; Conservative 0; Mismatches 2; Indels 2; Gaps 2;  
 QY 1 ATGAAACGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGCGCTCCCA 60

Db 175 ATGAAACGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACTGGCGCTCCCA 234  
 QY 61 GCAGCCACTCAGGAGAAAGATGTGCTGGGCAAAAAAGGGATACAGTGAACATGACC 120  
 Db 235 GCAGCCACTCAGGAGAAAGATGTGCTGGGCAAAAAAGGGATACAGTGAACATGACC 294  
 QY 121 TGTACGCTTCCAGAAAGAGACATCAATTCACCTGAAAAAACTCCAACAGATTAAG 180  
 Db 295 TGTACGCTTCCAGAAAGAGACATCAATTCACCTGAAAAAACTCCAACAGATTAAG 354  
 QY 181 ATTCTGGAAATCAGGGCTCTCTTAATAAAGGCTCAAGCTGAATATGCGGCT 240  
 Db 355 ATTCTGGAAATCAGGGCTCTCTTAATAAAGGCTCAAGCTGAATATGCGGCT 414  
 QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 300  
 Db 415 GACTCAAGAAAGACCTTTGGAGCAAGAAATTCCTCCCTGATCATCAAGATCTTAAG 474  
 QY 301 ATGAAAGCTCAGATATCTTACATCTGTGAATGTGAGAGACCAAGAGAGAGGTGCATTG 360  
 Db 475 ATGAAAGCTCAGATATCTTACATCTGTGAATGTGAGAGACCAAGAGAGAGGTGCATTG 534  
 QY 361 CTAGTGTGGATTGACTGCCAACTCTGACACCCACCTGCTCAGGGGCGAGAGCTTGAC 420  
 Db 535 CTAGTGTGGATTGACTGCCAACTCTGACACCCACCTGCTCAGGGGCGAGAGCTTGAC 594  
 QY 421 CTGACCTTGGAGAGCCCTCTGTTAGTACCCCTGATGATCAATGTAGAGTCCAAAGGGT 480  
 Db 595 CTGACCTTGGAGAGCCCTCTGTTAGTACCCCTGATGATCAATGTAGAGTCCAAAGGGT 654  
 QY 481 AAAAAATACA-GGGGGGAGAACCTCTCCGTCTCAGCTGAGAGCTCCAGATATGTGG 539  
 Db 655 AAAAAATACA-GGGGGGAGAACCTCTCCGTCTCAGCTGAGAGCTCCAGATATGTGG 714  
 QY 540 CACTGGAACATGACATCTCTTTCAGAAACCAAGAAAGGTGAGTCAAAATAGACATCGT 599  
 Db 715 CACTGGAACATGACATCTCTTTCAGAAACCAAGAAAGGTGAGTCAAAATAGACATCGT 774  
 QY 600 -GGTGCTAGCTTTCCAGAAAGGCTCC 624  
 Db 775 -GGTGCTAGCTTTCCAGAAAGGCTCC 800

RESULT 6  
 BI838893 840 bp mRNA linear EST 04-OCT-2001  
 LOCUS 603087279F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5226427 5',  
 DEFINITION mRNA sequence.  
 ACCESSION BI838893  
 VERSION BI838893.1 GI:15950443  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 AUTHORS 1 (bases 1 to 840)  
 JOURNAL NIH-MGC http://mgs.nci.nih.gov/  
 COMMENT National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Robert Strausberg, Ph.D.  
 Email: cgabbs-r@mail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LLNL at:  
 http://image.llnl.gov  
 Plate: LLAM1569 row: m column: 20  
 High quality sequence start: 11  
 High quality sequence stop: 832.

## FEATURES

Location/Qualifiers

source

1. 840  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5226427"  
/lab\_host="DH10B"  
/clone\_1lb="NIH\_MGC\_120"  
/note="Organ: pooled pancreas and spleen; Vector: pcwv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.9%; Score 589; DB 4; Length 840;  
Best Local Similarity 98.1%; Pred. No. 2e-143;  
Matches 611; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGGCTCTCCCA 60  
DB 168 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGGCTCTCCCA 227  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
DB 228 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 287  
QY 121 TGTACAGCTTCCGAAAGAAAGCATTAATTCCTGGAATACTCAACAGATTAAG 180  
DB 288 TGTACAGCTTCCGAAAGAAAGCATTAATTCCTGGAATACTCAACAGATTAAG 347  
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACTGAAGTCAATCAAGCTGAATGATCGGCT 240  
DB 348 ATTCTGGGAATCAGGGCTCTTCTTAACTGAAGTCAATCAAGCTGAATGATCGGCT 407  
QY 241 GATCAAGAAAGACCTTTGGGACCAAGAACTCCCGTGAATCAAGAAATCTTAAG 300  
DB 408 GATCAAGAAAGACCTTTGGGACCAAGAACTCCCGTGAATCAAGAAATCTTAAG 467  
QY 301 ATGAAGACTCAGATTAATCTATCTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
DB 468 ATGAAGACTCAGATTAATCTATCTGAAGTGAAGACCAAGAGAGAGTCAATTG 527  
QY 361 CTAGTGTTCGATTGACTGCCAATCTGCAACCACTGCTTCAAGGGCAGAGCTGACC 420  
DB 528 CTAGTGTTCGATTGACTGCCAATCTGCAACCACTGCTTCAAGGGCAGAGCTGACC 587  
QY 421 CTGACCTTGAAGAGCCCGCTGATGAGCCCTCAGTGAATGTAGAGTCCAAAGGGGT 480  
DB 588 CTGACCTTGAAGAGCCCGCTGATGAGCCCTCAGTGAATGTAGAGTCCAAAGGGGT 647  
QY 481 AAA-AAATATCAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATTAAGTGG 539  
DB 648 AAAAATATCAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATTAAGTGG 707  
QY 540 CACTGAGATGCACTGTCTTGGCAACCAAGAGAGTGAAGTCAAAAATGACATGCT 599  
DB 708 CACTGAGATGCACTGTCTTGGCAACCAAGAGAGTGAAGTCAAAAATGACATGCT 767  
QY 600 GG-TGCTAGCTTCCAGAGGCTCCCGGA 627  
DB 768 GGAGGCTAGCTTCCAGAGGCTCCCGCA 796

RESULT 7  
BI819727 791 bp mRNA linear EST 04-OCT-2001  
LOCUS 603041415F1 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5182274 5'  
DEFINITION mRNA sequence.  
ACCESSION BI819727

VERSION BI819727.1 GI:15931277  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
AUTHORS NIH-MGC <http://mgs.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: [cgapds-remail.nih.gov](mailto:cgapds-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LIML at: <http://image.llnl.gov>  
Plate: LIML1454 row: n column: 03  
High quality sequence stop: 782.  
Location/Qualifiers

FEATURES  
source  
1. 791  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5182274"  
/lab\_host="DH10B"  
/clone\_1lb="NIH\_MGC\_115"  
/note="Organ: pooled brain, lung, testis; Vector: pcwv-sport6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC Library."

ORIGIN

Query Match 40.1%; Score 577.4; DB 4; Length 791;  
Best Local Similarity 99.3%; Pred. No. 2.2e-140;  
Matches 611; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGGCTCTCCCA 60  
DB 176 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGGCTCTCCCA 235  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 120  
DB 236 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAATGACC 295  
QY 121 TGTACAGCTTCCGAAAGAAAGCATTAATTCCTGGAATACTCAACAGATTA-A 179  
DB 296 TGTACAGCTTCCGAAAGAAAGCATTAATTCCTGGAATACTCAACAGATTA-A 355  
QY 180 GATTCTGGGAATCAGGGCTCTTCTTAATAAGTCAATCCAGCTGAATGATGGCG 239  
DB 356 GATTCTGGGAATCAGGGCTCTTCTTAATAAGTCAATCCAGCTGAATGATGGCG 415  
QY 240 TGACTCAAGAAAGCTTTGGGACCAAGAACTCCCGTGAATCAAGAAATCTTA 299  
DB 416 TGACTCAAGAAAGCTTTGGGACCAAGAACTTTCCCTGATATCAAGAAATCTTA 475  
QY 300 GATGAAGAATCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGAGTGAATT 359  
DB 476 GATGAAGAATCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGAGTGAATT 535  
QY 360 GCTAGTGTTCGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGAC 419  
DB 536 GCTAGTGTTCGATTGACTGCACTGCAACCACTGCTTCAAGGGCAGAGCTGAC 595

QY 420 CTTGACCTTGG-AGAGCCCCCTGGTAGTAGACCCCTCAGTGCATGTAGGAGTCCAGGG 478  
 Db 596 CTTGACCTTGGCAGAGCCCCCTGGTAGTAGACCCCTCAGTGCATGTAGGAGTCCAGGG 655  
 QY 479 GTAAAAACATACAGGGGGGGAAGACCTCTCCGGTGTACAGCTGGAGCTCCAGGATATG 538  
 Db 656 GTAAAAACATACAGGGGGGGAAGACCTCTCCGGTGTACAGCTGGAGCTCCAGGATATG 715  
 QY 539 GCACCTGAGCA-TCGACTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATAGACATC 597  
 Db 716 GCACCTGAGCATTCGACTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATAGACATC 775  
 QY 598 GTGGTGTACCTTTC 612  
 Db 776 GTGGTGTACCTTTC 790

# RESULT 8

LOCUS  
 CV029573

DEFINITION  
 8403 Full Length cDNA from the Mammalian Gene Collection Homo  
 sapiens cDNA 5' similar to BC025782, mRNA sequence.  
 CV029573

VERSION  
 CV029573.1 GI:51487751

KEYWORDS  
 EST.  
 Homo sapiens (human)

SOURCE  
 ORGANISM

# REFERENCE

AUTHORS

1 (bases 1 to 557)  
 Rial,J.F., Hirozane-Kishikawa,T., Hao,T., Bertin,N., Li,S.,  
 Dicot,A., Li,N., Rosenberg,J., Lamesch,P., Vidalain,P.O.,  
 Clingingsmith,T.R., Hartley,J.L., Esposito,D., Cheo,D., Moore,T.,  
 Simmons,B., Segretera,R., Boeak,S., Doucette-Stamm,L., Le Peuch,C.,  
 Vandenhaute,J., Quisick,M.E., Alpala,J.S., Hill,D.E. and Vidal,M.  
 Human ORFeome Version 1.1: a Platform for Reverse Proteomics  
 Genome Res. (2004) In press

# TITLE

JOURNAL

COMMENT

Contact: Vidal M  
 Marc Vidal Laboratory  
 Dana Farber Cancer Institute  
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
 Tel: 617 632 5180  
 Fax: 617 632 5739  
 Email: Marc.Vidal@dfci.harvard.edu  
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF  
 results from a PCR reaction using an MGC full-length cDNA as  
 template DNA and ORF specific primers  
 PCR Primers  
 FORWARD: ATGACCGGGAGTCCCT  
 BACKWARD: CAATGGGGTACATGCTTCTGAAA  
 Insert Length: 557 Std Error: 21.00  
 Plate: 11064 row: 03 column: B  
 Seq primer: ACTGGCGGCTTTTACACGCTGCTGACTGGAAAC  
 High quality sequence start: 98  
 High quality sequence stop: 556  
 POLYA=No.

# FEATURES

source

1..557  
 /organism="Homo sapiens"  
 /mol\_type="RNA"  
 /db\_xref="taxon:9606"  
 /tissue\_type="mixed"  
 /clone\_lib="Full Length cDNA from the Mammalian Gene  
 Collection"  
 /note="Vector: mixed. The ORFs were PCR amplified from the  
 MGC (Mammalian Gene Collection) as of April 2004 and  
 cloned by recombinational Gateway cloning into pDONR223  
 Donor vector. Reference : MGC (Mammalian Gene Collection)  
 Program Team, Generation and Initial Analysis of more than  
 15,000 Full-length Human and Mouse cDNA Sequences. PNAS,  
 2002, 99(26), 16899-16903"

# ORIGIN

Query Match 38.6%; Score 555.4; DB 7; Length 557;  
 Best Local Similarity 99.8%; Pred. No. 1.2e-134;  
 Matches 556; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTTGTGCTGTCACACTGGCGCTCCGCCA 60  
 Db 1 ATGAACCGGGAGTCCCTTTAGGCACTGCTTGTGCTGTCACACTGGCGCTCCGCCA 60  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 Db 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 QY 121 TGTACGCTTCCGAGAAGACATCAATTCCATGAAAAAATCTCAACAGATTAAG 180  
 Db 121 TGTACGCTTCCGAGAAGACATCAATTCCATGAAAAAATCTCAACAGATTAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCCAGGCTGAATGCGGCT 240  
 Db 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCCAGGCTGAATGCGGCT 240  
 QY 241 GACTCAAGAAGAGCTTTGGAGCCAGAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 Db 241 GACTCAAGAAGAGCTTTGGAGCCAGAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 QY 301 ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 360  
 Db 301 ATGAAAGACTCAGATATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 360  
 QY 361 CTAGTGTGGATTGCTGCTGCAACTGTGACCACTGCTTCAAGGGCAGAGCTGACC 420  
 Db 361 CTAGTGTGGATTGCTGCTGCAACTGTGACCACTGCTTCAAGGGCAGAGCTGACC 420  
 QY 421 CTGACCTTGAAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGTCAAGGGT 480  
 Db 421 CTGACCTTGAAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGTCAAGGGT 480  
 QY 481 AAAAATATCAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGAGACTCCAGATATGTGC 540  
 Db 481 AAAAATATCAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGAGACTCCAGATATGTGC 540  
 QY 541 ACCTGAGCATGCACTGT 557  
 Db 541 ACCTGAGCATGCACTGT 557

# RESULT 9

LOCUS  
 BX437619

DEFINITION  
 BX437619 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP007YD02  
 5-PRIME, mRNA sequence.  
 BX437619

ACCESSION  
 BX437619.2 GI:4699918

VERSION  
 EST.  
 SOURCE  
 ORGANISM

# REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 713)  
 Li,W.B., Gruber,C., Jesse,J. and Polayes,D.  
 Full-length cDNA libraries and normalization  
 unpublished (2001)  
 On May 15, 2003 this sequence version replaced gi:30771588.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
 Email: sequefgenoscope.cns.fr Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see

FEATURES  
source

http://www.genoscope.cns.fr/cdnas=CS0CAP007DB010P1ec=6485.r.  
Location/Qualifiers  
1. 713  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CS0CAP007YD02"  
/issue\_type="THYMUS"  
/clone\_lib="Homo sapiens THYMUS"  
/note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."

## ORIGIN

Query Match 37.4%; Score 539; DB 5; Length 713;  
Best Local Similarity 99.5%; Pred. No. 2,66-130;  
Matches 550; Conservative 1; Mismatches 1; Indels 1; Gaps 1;

```

QY 1 ATGAACGGGGAGTCCCTTTTAGGACATGCTTCTGTGGCTGCAACTGGGCTCTCCCA 60
Db 161 ATGAACCGGGAGTCCCTTTTAGGACATGCTTCTGTGGCTGCAACTGGGCTCTCCCA 220
QY 61 GCAGCCACTGAGGGAAGAAAGTGTCTGGGCAAAAAAGGGGATAGAGTGAATGACC 120
Db 221 GCAGCCACTGAGGGAAGAAAGTGTCTGGGCAAAAAAGGGGATAGAGTGAATGACC 280
QY 121 TGTAAGCTTCCGAAAGAGATCAATTCACCTGGAAGAACTCCAAACAGATTAAG 180
Db 281 TGTAAGCTTCCGAAAGAGATCAATTCACCTGGAAGAACTCCAAACAGATTAAG 340
QY 181 ATTCTGGGAATCGGGCTCTCTTAACCTAAGCTCATCCAGCTGAATGAGGGCT 240
Db 341 ATTCTGGGAATCGGGCTCTCTTAACCTAAGCTCATCCAGCTGAATGAGGGCT 400
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 401 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 460
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360
Db 461 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 520
QY 361 CTAGTGTTCGATGATGCTGCAACTGACCACTGCTTCAAGGGGCAAGGCTAACC 420
Db 521 CTAGTGTTCGATGATGCTGCAACTGACCACTGCTTCAAGGGGCAAGGCTAACC 580
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 480
Db 581 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGGGT 640
QY 481 AAAAACAATACA-GGGGGGGAAGACCTCTCCGTGTCTCACTGAGCTCCAGATAGTG 539
Db 641 AAAAACAATACAGGGGGGGAAGACCTCTCCGTGTCTCACTGAGCTCCAGATAGTG 700
QY 540 CACCTGACATGC 552
Db 701 CACCTGACATGC 713

```

RESULT 10  
LOCUS B1915265 723 bp mRNA linear EST 16-OCT-2001  
DEFINITION 60318431F1 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5248226 5',  
mRNA sequence.  
ACCESSION B1915265  
VERSION B1915265.1 GI:16179363  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Cranialia; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 723)  
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cga@gsf.com, r-straus@nci.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
DNA Sequencing by: Invitrogen Genomics, Inc.  
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LINL at:  
http://image.lnl.gov  
Plate: L14M1626 row: j column: 03  
High quality sequence stop: 717.

## FEATURES

## source

Location/Qualifiers  
1. 723  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5248226"  
/lab\_host="DH10B"  
/clone\_lib="NIH\_MGC\_121"  
/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.7 kb, insert size range 0.7-3.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 017. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 36.1%; Score 519.4; DB 4; Length 723;  
Best Local Similarity 99.1%; Pred. No. 3,8e-125;  
Matches 564; Conservative 0; Mismatches 1; Indels 4; Gaps 4;

```

QY 1 ATGAACGGGGAGTCCCTTTTAGGACATGCTTCTGTGGCTGCAACTGGGCTCTCCCA 60
Db 157 ATGAACCGGGAGTCCCTTTTAGGACATGCTTCTGTGGCTGCAACTGGGCTCTCCCA 216
QY 61 GCAGCCACTGAGGGAAGAAAGTGTCTGGGCAAAAAAGGGGATAGAGTGAATGACC 119
Db 217 GCAGCCACTGAGGGAAGAAAGTGTCTGGGCAAAAAAGGGGATAGAGTGAATGACC 276
QY 120 CTGACAGCTTCCGAAAGAGATCAATTCACCTGGAAGAACTCCAAACAGATTAAG 179
Db 277 CTGACAGCTTCCGAAAGAGATCAATTCACCTGGAAGAACTCCAAACAGATTAAG 336
QY 180 GATTCTGGGAATGAGGGCTCTTTTAAGTGAAGTGCATCCAGTGAATGATGCCG 239
Db 337 GATTCTGGGAATGAGGGCTCTTTTAAGTGAAGTGCATCCAGTGAATGATGCCG 396
QY 240 TGACTCAAGAAAGCCTTTGGGACCAAGGAACCTCCCTGATCATCAAGATCTTA 299
Db 397 TGACTCAAGAAAGCCTTTGGGACCAAGGAACCTCCCTGATCATCAAGATCTTA 456
QY 300 GATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATT 359
Db 457 GATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATT 516
QY 360 GCTAGTGTGATGATGATGCTGCAACTGACCACTGCTTCAAGGGGA-GAGCTGA 418
Db 517 GCTAGTGTGATGATGATGCTGCAACTGACCACTGCTTCAAGGGGATGAGCTGA 576
QY 419 CCTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGG 478
Db 577 CCTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGAATGTAGAGATCCAAAGG 636
QY 479 GTAAAAACATACAGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTG 538

```

Db 637 GTAAAAACATACA-GGGGGGAAGACCTCTCCGTCTCTCACTGAGCTCCAGATAGTG 635

QY 539 GCACCTGACATGACATGCTCTTGAGAAC 567

Db 696 G-ACCTGGACATGACATGCTCTTGAGAAC 723

RESULT 11  
CD609191 679 bp mRNA linear EST 12-JAN-2004  
LOCUS 5606993JH1 FLP Homo sapiens cDNA, mRNA sequence.  
DEFINITION CD609191  
ACCESSION CD609191  
VERSION CD609191.1 GI:40257454  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 679)  
AUTHORS Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.  
TITLE Circular rapid amplification of cDNA ends for high-throughput extension cloning of partial genes  
JOURNAL Genomics 84 (1), 205-210 (2004)  
COMMENT Contact: Fu GK  
Incyte Genomics, Inc.  
3160 Porter Dr., Palo Alto, CA 94304, USA  
Tel: 6508454102  
Email: gfu@incyte.com.

FEATURES  
source 1..679  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="FLP"  
/note="Vector: pDrive Cloning Vector"

ORIGIN

Query Match 35.2%; Score 506.4; DB 6; Length 679;  
Best Local Similarity 99.8%; Pred. No. 9,7e-122;  
Matches 507; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACGGGCTCCCTCCA 60  
Db 172 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACGGGCTCCCTCCA 231

QY 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAAGGGATACAGTGAATGACC 120  
Db 232 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAAGGGATACAGTGAATGACC 231

QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAAACTCCACACAGATTAAG 180  
Db 292 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAAACTCCACACAGATTAAG 351

QY 181 ATTCTGGGAATCAGGGCTCTCTTTTAACATTAAGTCCATCCAGCTGAATGATCGGCT 240  
Db 352 ATTCTGGGAATCAGGGCTCTCTTTTAACATTAAGTCCATCCAGCTGAATGATCGGCT 411

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
Db 412 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 471

QY 301 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
Db 472 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 531

QY 361 CTGATGTTGGATTAATGCACTGCACTGCAACCACTGCTTGAAGGGGACAGCTGACC 420  
Db 532 CTGATGTTGGATTAATGCACTGCAACCACTGCTTGAAGGGGACAGCTGACC 591

QY 421 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCATGATGATGATGATGATG 480  
Db 592 CTGACCTTGGAGAGCCCTCTGTGATGAGCCCTCAGTGCATGATGATGATGATGATG 651

QY 481 AAAAACATACAGGGGGGAAGACCTCT 508

Db 652 AAAAACATACAGGGGGGAAGACCTCT 679

RESULT 12  
AUI41298 732 bp mRNA linear EST 05-AUG-2002  
LOCUS AUI41298 THYR01 Homo sapiens cDNA clone THYR01000359 5', mRNA  
DEFINITION AUI41298  
ACCESSION AUI41298  
VERSION AUI41298.1 GI:11002819  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE 1 (bases 1 to 732)  
AUTHORS Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saico,K., Kawai,Y., Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and Isogai,T.  
TITLE HRI human cDNA project  
JOURNAL Unpublished (2000)  
COMMENT Contact: Takao Isogai  
Genomics Laboratory  
Helix Research Institute  
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan  
Tel: 81-438-52-3975  
Fax: 81-438-52-3986  
Email: genomics@hri.co.jp  
HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix Research Institute; cDNA library construction: Department of Virology, Institute of Medical Science, University of Tokyo, and Helix Research Institute.

FEATURES  
source 1..732  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone\_lib="THYR01000359"  
/issue\_type="Chryoid gland"  
/clone\_lib="THYR01"  
/note="Vector: pME18SFL3"

ORIGIN

Query Match 35.0%; Score 504; DB 1; Length 732;  
Best Local Similarity 97.2%; Pred. No. 4.2e-121;  
Matches 522; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACCTGCGCTCCTCCA 60  
Db 196 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGCTGCAACCTGCGCTCCTCCA 255

QY 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAAGGGATACAGTGAATGACC 120  
Db 256 GCAGCCACTCAGGGAAGAAAGTGCTGCTGGCAAAAAAGGGATACAGTGAATGACC 315

QY 121 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAAACTCCACACAGATTAAG 180  
Db 316 TGTACAGCTTCCAGAAAGAGCATCAATTCACATGAAAACTCCACACAGATTAAG 375

QY 181 ATTCTGGGAATCAGGGCTCTCTTTTAACATTAAGTCCATCCAGCTGAATGATCGGCT 240  
Db 376 ATTCTGGGAATCAGGGCTCTCTTTTAACATTAAGTCCATCCAGCTGAATGATCGGCT 435

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
Db 436 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 495

QY 301 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
Db 496 ATAGAACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 555

QY 361 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGAGAGCCTGACC 420  
 DB 556 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGAGAGCCTGACC 615  
 QY 421 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
 DB 616 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 675  
 QY 481 AAAAACAATACAGGGGGGGG--AAGACCTCTCCGCTGCTGAGCTGAGACTCCAGGATA 535  
 DB 676 AAAAACAATACAGGGGGGGGAAACCTTTTCGGGGCTTAACCTTGAGCTTCANAGATA 732

RESULT 13  
 CO246446 754 bp mRNA linear EST 23-JUN-2004  
 LOCUS DEFINITION AGENCOURT\_26517764 NIH\_MGC\_212 Homo sapiens cDNA clone  
 CO246446 IMAGE:30924065 5', mRNA sequence.  
 CO246446.1 GI:49111174

EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 754)  
 NIH-MGC http://mgs.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892

COMMENT  
 Email: cgabs-remail.nih.gov  
 Tissue Procurement: Mary Hendrix  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: NDAMI165 row: 0 column: 18  
 High quality sequence stop: 507.

FEATURES  
 source location/Qualifiers

1..754  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30924065"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="NIH\_MGC\_212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site: 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bernaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN  
 Query Match 33.9%; Score 488.8; DB 7; Length 754;  
 Best Local Similarity 96.6%; Pred. No. 4.2e-117;  
 Matches 510; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTGAAGCACTTGTGCTGCTGCAACTGGCGCTCTCCCA 60  
 DB 173 ATGAACCGGGAGTCCCTTTTGAAGCACTTGTGCTGCTGCAACTGGCGCTCTCCCA 232

QY 61 GCAGCCACTCAGGAGAAAGAAAGTGTCTGTGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 233 GCAGCCACTCAGGAGAAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 292  
 QY 121 TGTACAGCTTCCCAAGAAAGAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 180  
 DB 293 TGTACAGCTTCCCAAGAAAGAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 352  
 QY 181 ATTCTGGGAAATAGAGGCTCTCTTAACTAAAGATCCATCCAGGCTGATGCGGCT 240  
 DB 353 ATTCTGGGAAATAGAGGCTCTCTTAACTAAAGATCCATCCAGGCTGATGCGGCT 412  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCTGATCATCAAGATCTTAAG 300  
 DB 413 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCTGATCATCAAGATCTTAAG 472  
 QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGGAGACCAAGAGAGGTGCAATTG 360  
 DB 473 ATGAAGACTCAGATACCTTACATCTGTGAAGTGGAGACCAAGAGAGGTGCAATTG 532  
 QY 361 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGAGAGCCTGACC 420  
 DB 533 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACCTGCTTCAGGGGAGAGCCTGACC 592  
 QY 421 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 480  
 DB 593 CTGACCTTGGAGAGCCCCCTGCTGATGAGCCCTCACTGCAATGTAGAGTCCAGGGGT 652  
 QY 481 AAAAACAATACAGGGGGGAAAGACCTTCTCGTCTCTGAGCTGAGCTC 528  
 DB 653 AAAAACAATACAGGGGGGAAAGACCTTCTCGTCTCTGAGCTGAGCTC 699

RESULT 14  
 CF125444 710 bp mRNA linear EST 05-AUG-2003  
 LOCUS DEFINITION UI-HF-E10-avo-e-23-0-UI.r1 NIH\_MGC\_212 Homo sapiens cDNA clone  
 CF125444 IMAGE:30563734 5', mRNA sequence.  
 CF125444.1 GI:33201679

EST.  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

REFERENCE 1 (bases 1 to 710)  
 Bernaldo, M.F., Lennon, G. and Soares, M.B.  
 Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 Genome Res. 6 (9), 791-806 (1996)

JOURNAL MEDLINE  
 PUBMED  
 COMMENT  
 Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu

Tissue Procurement: Mary Hendrix  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 Seq primer: pYX-5.

FEATURES  
 source location/Qualifiers

1..710  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30563734"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"

/lab\_host="DH10B (T1 phage resistant)"  
 /clone\_11b="NIH MG-212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bernaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGGCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.9%; Score 488.4; DB 7; Length 710;  
 Best Local Similarity 98.9%; Pred. No. 5.3e-117;  
 Matches 522; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGTGTGTGCAACTGGCGCTCTCCCA 60  
 DB 186 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGTGTGTGCAACTGGCGCTCTCCCA 245  
 QY 61 GCAGCCACTAGAGGAAAGAAAGTGTCTGTGGCAAAAAGGGATACAGTGAACCTGACC 120  
 DB 246 GCAGCCACTAGAGGAAAGAAAGTGTCTGTGGCAAAAAGGGATACAGTGAACCTGACC 305  
 QY 121 TGTACAGCTTCCAGAAAGAGATATTCATTCACAGTGAATACTCCACAGATTAAG 180  
 DB 306 TGTACAGCTTCCAGAAAGAGATATTCATTCACAGTGAATACTCCACAGATTAAG 365  
 QY 181 ATTCTGGGAATGAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 240  
 DB 366 ATTCTGGGAATGAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 425  
 QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTAAG 300  
 DB 426 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTAAG 485  
 QY 301 ATGAAGACTCAGATATCTACATCTGTGAAGTGAAGAGCAGAAAGAGAGTGCATTTG 360  
 DB 486 ATGAAGACTCAGATATCTACATCTGTGAAGTGAAGAGCAGAAAGAGTGCATTTG 545  
 QY 361 CTAGTGTTCGATTAAGTTCGCAACTGTGACACCACTGTCTTCAAGGGGACGCTGACC 420  
 DB 546 CTAGTGTTCGATTAAGTTCGCAACTGTGACACCACTGTCTTCAAGGGGACGCTGACC 605  
 QY 421 CTGACCTTGAAGAGCCCTTGTGTAGAGCCCTCACTGCAATGTAGAGATCCAAAGGGCT 480  
 DB 606 CTGACCTTGAAGAG-CCCTCTGTGTAGAG-CCCTCACTGCAATGTAGAGATCCAAAGGGCT 663  
 QY 481 AAAAACAATACAGGGGGAGAGACCTCTCGTGTCTGAGCTGAGGCTC 528  
 DB 664 AAAAACAATACAGGGGGAGAGACCTCTCGTGTCTGAGCTGAGGCTC 710

RESULT 15  
 LOCUS CB052686 694 bp mRNA linear EST 17-JUN-2003  
 DEFINITION NISC\_gli9e06.y1 NCI\_CGAP\_Le12 Homo sapiens cDNA clone IMAGE:3290578  
 5', mRNA sequence.  
 ACCESSION CB052686  
 VERSION CB052686.1 GI:27790973  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)

## COMMENT

Contact: Robert Strausberg, Ph.D.  
 Email: cgaabs-r@mail.nih.gov  
 CDNA Library Preparation:  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
 DNA Sequencing by: National Institutes of Health Intramural  
 Sequencing Center (NISC)  
 Clone distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 info@image.llnl.gov  
 Plate: LLAM8058 row: I column: 11  
 Seq primer: M13Rpi reverse primer (ABI).  
 Location/Qualifiers  
 1. 694

## FEATURES

Source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3290578"  
 /issue\_type="leiomysarcoma"  
 /lab\_host="DH10B"  
 /clone\_11b="NCI\_CGAP\_Le12"  
 /note="Organ: soft tissue; Vector: pT7T3D-Pac (Pharmacia)  
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5'-ACTGGAAGATTCGGCGCGCATCTTTTCTTTTCTTTTCTTTT-3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pT7T3 vector. Library  
 went through one round of normalization. Library  
 constructed by Bento Soares and M. Fatima Bernaldo."

## ORIGIN

Query Match 33.6%; Score 483.8; DB 6; Length 694;  
 Best Local Similarity 99.4%; Pred. No. 8.5e-116;  
 Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGTGTGTGCAACTGGCGCTCTCCCA 60  
 DB 196 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGTGTGTGCAACTGGCGCTCTCCCA 255  
 QY 61 GCAGCCACTCAGGAAAGAAAGTGTGTGTGGC-AAAAAGGGATACAGTGAACCTGAC 119  
 DB 256 GCAGCCACTCAGGAAAGAAAGTGTGTGTGGC-AAAAAGGGATACAGTGAACCTGAC 315  
 QY 120 CTGTACAGCTTCCAGAAAGAGATATTCATTCAGTGAATACTCCACAGATTA 179  
 DB 316 CTGTACAGCTTCCAGAAAGAGATATTCATTCAGTGAATACTCCACAGATTA 375  
 QY 180 GATTCGGGAATTCAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 239  
 DB 376 GATTCGGGAATTCAGGGCTCTCTTAACTAAAGTTCATCAAGCTGAATGATCGGCT 435  
 QY 240 TGACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTA 299  
 DB 436 TGACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTA 495  
 QY 240 TGACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTA 299  
 DB 436 TGACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGTATCATCAAGATTTTA 495  
 QY 300 GATGAAGAAGCTGATTAATCTGTGAAGTGAAGAGCAGAAAGAGAGGTGCAATT 359  
 DB 496 GATGAAGAAGCTGATTAATCTGTGAAGTGAAGAGCAGAAAGAGAGGTGCAATT 555  
 QY 360 GCTAGTTCGATTAAGTTCGCAACTGTGACACCACTGTCTTCAAGGGGACAGCTGAC 419  
 DB 556 GCTAGTTCGATTAAGTTCGCAACTGTGACACCACTGTCTTCAAGGGGACAGCTGAC 615  
 QY 420 CCTGACCTTGAAGAGCCCTTGTGTAGTGAAGCCCTCAGTGAATGAAGTCCAAAGGG 479  
 DB 616 CCTGACCTTGAAGAGCCCTTGTGTAGTGAAGCCCTCAGTGAATGAAGTCCAAAGGG 675  
 QY 480 TAAAAACATACAGGGGGG 498  
 DB 676 TAAAAACATACAGGGGGG 694

Search completed: September 20, 2005, 12:03:00

Wed Sep 28 10:02:31 2005

us-09-936-702-4.rst

Page 11

Job time : 5149 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 26, 2005, 13:01:41 / Search time 161 Seconds  
(without alignments)  
622.180 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPKGSSVVKY.....RYYFGQGTRLIKLVERGSG 259

Scoring table: OLIGO

Gapop 60.0, Gapext 60.0

Searched: 2105692 seqs, 386760381 residues

Word size: 0

Total number of hits satisfying chosen parameters: 2815

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database: A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	21	8.1	259	5	ABP45023 Human Bly
2	21	8.1	259	7	ADG95850 Single ch
3	20	7.7	259	5	ABP45252 Human Bly
4	20	7.7	259	5	ABP45264 Human Bly
5	20	7.7	259	5	ABP45236 Human Bly
6	20	7.7	259	5	ABP45032 Human Bly
7	20	7.7	259	5	ABP45259 Human Bly
8	20	7.7	259	5	ABP45486 Human Bly
9	20	7.7	259	7	ADG96086 Single ch
10	20	7.7	259	7	ADG96091 Single ch
11	20	7.7	259	7	ADG96313 Single ch
12	20	7.7	259	7	ADG96063 Single ch
13	20	7.7	259	7	ADG95859 Single ch
14	20	7.7	259	7	ADG96079 Single ch
15	18	6.9	259	5	AU72862 P4-3 sing
16	16	6.2	259	3	AAB09775 TWY 30K m
17	16	6.2	259	5	ABP45345 Human Bly
18	16	6.2	259	5	ABP44853 Human Bly
19	16	6.2	259	5	ABP45474 Human Bly
20	16	6.2	259	5	ABP45441 Human Bly
21	16	6.2	259	5	ABP45470 Human Bly
22	16	6.2	259	5	ABP45021 Human Bly
23	16	6.2	259	5	ABP45030 Human Bly
24	16	6.2	259	5	ABP45556 Human Bly
25	16	6.2	259	5	ABP45014 Human Bly

26	16	6.2	259	5	ABP45653 Human Bly
27	16	6.2	259	5	ABP45266 Human Bly
28	16	6.2	259	5	ABP45541 Human Bly
29	16	6.2	259	5	ABP44962 Human Bly
30	16	6.2	259	5	ABP45251 Human Bly
31	16	6.2	259	7	ADG32324 Mouse scf
32	16	6.2	259	7	ADG32322 Mouse scf
33	16	6.2	259	7	ADG35857 Single ch
34	16	6.2	259	7	ADG96383 Single ch
35	16	6.2	259	7	ADG96368 Single ch
36	16	6.2	259	7	ADG96480 Single ch
37	16	6.2	259	7	ADG96268 Single ch
38	16	6.2	259	7	ADG95680 Single ch
39	16	6.2	259	7	ADG96301 Single ch
40	16	6.2	259	7	ADG96093 Single ch
41	16	6.2	259	7	ADG96297 Single ch
42	16	6.2	259	7	ADG96172 Single ch
43	16	6.2	259	7	ADG95841 Single ch
44	16	6.2	259	7	ADG96078 Single ch
45	16	6.2	259	7	ADG95789 Single ch

## ALIGNMENTS

RESULT 1  
ABP45023  
ID ABP45023 standard; protein; 259 AA.

AC ABP45023;  
DT 19-AUG-2002 (first entry)  
XX Human Blys binding scfv seq ID 1034.

XX Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
XX tumour necrosis factor; B cell proliferation; B cell differentiation;  
XX immunosuppressive; immunostimulant; antirheumatic;  
XX antiAids; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
XX common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.  
XX OS  
XX PN WO200202641-A1.  
XX PD 10-JAN-2002.

XX 15-JUN-2001; 2001WO-US019110.  
XX 16-JUN-2000; 2000US-0212210P.  
XX 17-OCT-2000; 2000US-0240816P.  
XX 16-MAR-2001; 2001US-0276248P.  
XX 21-MAR-2001; 2001US-0277379P.  
XX 25-MAY-2001; 2001US-0293499P.

XX (HUMA-) HUMAN GENOME SCI INC.  
XX (CAME-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for the  
XX diagnosis and treatment of cancers and immune disorders.

XX Claim 1; Page 1635-1636; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to  
XX B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
XX tumour necrosis factor (TNF) super family and induces B cell  
XX proliferation and differentiation. The antibodies of the invention have  
XX cytostatic, immunosuppressive, immunostimulant, immunomodulatory,

CC antineumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP4990-ABP47228 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention

XX SQ Sequence 259 AA;

Query Match 8.1%; Score 21; DB 5; Length 259;

Best Local Similarity 100.0%; Pred. No. 3.7e-10; Mismatches 0; Indels 0; Gaps 0;

XX DB 6 ESGAEVKKPQSSVVKVSCKASG 26  
 6 ESGAEVKKPQSSVVKVSCKASG 26

RESULT 2  
 ADG95850  
 ID ADG95850 standard; protein; 259 AA.

XX AC ADG95850;

XX DT 11-MAR-2004 (first entry)

XX DE Single chain antibody that immunospecifically binds Blys SEQID 1034.

XX KM Antibody; B lymphocyte stimulator; Blys; tumor necrosis factor;  
 KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
 KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukemia;  
 KM carcinoma; lymphoma; antineumatic; antiarthritic; neuroprotective;  
 KM antiinflammatory; antiaesthetic; antiallergic; cytostatic.

XX OS Unidentified.

XX PN WO2003055979-A2.

XX PD 10-JUL-2003.

XX PF 14-NOV-2002; 2002WO-US036496.

XX PR 16-NOV-2001; 2001US-0331469P.

XX PR 19-DEC-2001; 2001US-0340817P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX DR WPI; 2003-505530/47.

XX PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
 PT (Blys), useful for detecting and treating diseases or disorders e.g.  
 PT rheumatoid arthritis, asthma and leukemia.

XX PS Example 1; SEQ ID NO 1034; 394pp; English.

XX CC This invention relates to novel antibodies that immunospecifically bind  
 CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
 CC chromosome 13q34 and encodes a protein that is a member of the tumor  
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
 CC proliferation and differentiation. Specifically, it refers to single  
 CC chain antibody molecules (scfvs) derived, preferably, from the variable  
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
 CC fragment thereof, of either human, murine, rat or monkey Blys. The  
 CC present invention refers to the use of such antibodies in various methods  
 CC for the detection, diagnosis and prognosis of diseases related to the

CC aberrant expression or inappropriate function of Blys or its receptor. As  
 CC such, these compositions are useful for identifying immune disorders  
 CC including myasthenia gravis and multiple sclerosis, inflammatory  
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
 CC as AIDS and proliferative disorders including leukemia, carcinoma and  
 CC lymphoma. Accordingly, they can be described as exhibiting various  
 CC activities such as antineumatic, antiarthritic, neuroprotective,  
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX SQ Sequence 259 AA;

Query Match 8.1%; Score 21; DB 7; Length 259;

Best Local Similarity 100.0%; Pred. No. 3.7e-10; Mismatches 0; Indels 0; Gaps 0;

XX DB 6 ESGAEVKKPQSSVVKVSCKASG 26  
 6 ESGAEVKKPQSSVVKVSCKASG 26

RESULT 3  
 ABP45252  
 ID ABP45252 standard; protein; 259 AA.

XX AC ABP45252;

XX DT 19-ANG-2002 (first entry)

XX DE Human Blys binding scfv SEQ ID 1263.

XX KM Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KM tumor necrosis factor; B cell proliferation; B cell differentiation;  
 KM immunosuppressive; immunostimulant; immunomodulatory; antineumatic;  
 KM antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KM systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KM common variable immunodeficiency; acquired immunodeficiency syndrome.

XX OS Homo sapiens.

XX PN WO200202641-A1.

XX PD 10-JAN-2002.

XX PF 15-JUN-2001; 2001WO-US019110.

XX PR 16-JUN-2000; 2000US-0212210P.

XX PR 17-OCT-2000; 2000US-0240816P.

XX PR 16-MAR-2001; 2001US-0276248P.

XX PR 21-MAR-2001; 2001US-0277379P.

XX PR 25-MAY-2001; 2001US-0293499P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PA (CAMP-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX DR WPI; 2002-114799/15.

XX PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.

XX PS Claim 1; Page 1909-1911; 3148pp; English.

XX CC This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumor necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antineumatic and antiAIDS activity and can be used in vaccines to

CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP4728 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention  
 CC  
 SQ Sequence 259 AA;  
 QY  
 Db 7 SGAEEVKKPGSSVYVSCKASG 26  
 7 SGAEEVKKPGSSVYVSCKASG 26  
 RESULT 4  
 ID ABP45264 standard; protein; 259 AA.  
 AC ABP45264;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human Blys binding scFv SEQ ID 1275.  
 XX  
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200202641-A1.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001WO-US019110.  
 XX  
 PR 16-JUN-2000; 2000US-0212210P.  
 PR 17-OCT-2000; 2000US-0240816P.  
 PR 16-MAR-2001; 2001US-0276248P.  
 PR 21-MAR-2001; 2001US-0277379P.  
 PR 25-MAY-2001; 2001US-0293499P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 XX  
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
 XX  
 DR WPI; 2002-114799/15.  
 XX  
 PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 XX  
 PS Claim 1; Page 1924-1925; 3148pp; English.  
 XX  
 CC This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumour necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antirheumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys

CC and so may be used to detect and quantitate the presence of Blys in  
 CC biological samples and may be used in this way to diagnose disease  
 CC associated with aberrant expression of Blys. They may also be  
 CC administered to treat diseases associated with aberrant Blys expression  
 CC and activity such as cancer, immune, and autoimmune disorders and  
 CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
 CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
 CC acquired immunodeficiency syndrome (AIDS)). ABP4390-ABP4728 represent  
 CC the antibodies and fragments of the antibodies described in the method of  
 CC the invention  
 CC  
 SQ Sequence 259 AA;  
 QY  
 Db 7 SGAEEVKKPGSSVYVSCKASG 26  
 7 SGAEEVKKPGSSVYVSCKASG 26  
 RESULT 5  
 ID ABP45236 standard; protein; 259 AA.  
 AC ABP45236;  
 XX  
 DT 19-AUG-2002 (first entry)  
 XX  
 DE Human Blys binding scFv SEQ ID 1247.  
 XX  
 KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
 KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
 KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
 KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
 KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
 KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200202641-A1.  
 XX  
 PD 10-JAN-2002.  
 XX  
 PF 15-JUN-2001; 2001WO-US019110.  
 XX  
 PR 16-JUN-2000; 2000US-0212210P.  
 PR 17-OCT-2000; 2000US-0240816P.  
 PR 16-MAR-2001; 2001US-0276248P.  
 PR 21-MAR-2001; 2001US-0277379P.  
 PR 25-MAY-2001; 2001US-0293499P.  
 XX  
 PA (HUMA-) HUMAN GENOME SCI INC.  
 PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
 XX  
 PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
 XX  
 DR WPI; 2002-114799/15.  
 XX  
 PT Antibodies against B lymphocyte stimulating polypeptides, useful for the  
 PT diagnosis and treatment of cancers and immune disorders.  
 XX  
 PS Claim 1; Page 1890-1891; 3148pp; English.  
 XX  
 CC This invention describes novel antibodies that immunospecifically bind to  
 CC B lymphocyte stimulator (Blys) polypeptides. Blys is a member of the  
 CC tumour necrosis factor (TNF) super family and induces B cell  
 CC proliferation and differentiation. The antibodies of the invention have  
 CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
 CC antirheumatic and antiAIDS activity and can be used in vaccines to  
 CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
 CC and so may be used to detect and quantitate the presence of Blys in

CC biological samples and may be used in this way to diagnose disease  
CC associated with aberrant expression of Blys. They may also be  
CC administered to treat diseases associated with aberrant Blys expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP4728 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26

RESULT 6  
ID ABP45032 standard; protein; 259 AA.  
XX  
AC ABP45032;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human Blys binding scFv SEQ ID 1043.  
XX  
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX  
OS Homo sapiens.  
XX  
PN MO200202641-A1.  
XX  
PD 10-JAN-2002.  
XX  
PE 15-JUN-2001; 2001MO-US019110.  
XX  
PR 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX  
DR WPI; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte Stimulating polypeptides, useful for the  
PT diagnosis and treatment of cancers and immune disorders.  
XX  
Claim 1, Page 1646-1648; 3148pp; English.  
  
This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
CC antirheumatic and antiAIDS activity and can be used in vaccines to  
CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
CC and so may be used to detect and quantitate the presence of Blys in  
CC biological samples and may be used in this way to diagnose disease

CC associated with aberrant expression of Blys. They may also be  
CC administered to treat diseases associated with aberrant Blys expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP4728 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention  
XX  
SQ Sequence 259 AA;  
  
Query Match 7.7%; Score 20; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
Qy 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26

RESULT 7  
ID ABP45259 standard; protein; 259 AA.  
XX  
AC ABP45259;  
XX  
DT 19-AUG-2002 (first entry)  
XX  
DE Human Blys binding scFv SEQ ID 1270.  
XX  
KW Blys; B lymphocyte stimulator; TNF superfamily; human; cytostatic;  
KW tumour necrosis factor; B cell proliferation; B cell differentiation;  
KW immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
KW antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
KW systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
KW common variable immunodeficiency; acquired immunodeficiency syndrome.  
XX  
OS Homo sapiens.  
XX  
PN MO200202641-A1.  
XX  
PD 10-JAN-2002.  
XX  
PE 15-JUN-2001; 2001MO-US019110.  
XX  
PR 16-JUN-2000; 2000US-0212210P.  
PR 17-OCT-2000; 2000US-0240816P.  
PR 16-MAR-2001; 2001US-0276248P.  
PR 21-MAR-2001; 2001US-0277379P.  
PR 25-MAY-2001; 2001US-0293499P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
PA (CAMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;  
XX  
DR WPI; 2002-114799/15.  
XX  
PT Antibodies against B lymphocyte Stimulating polypeptides, useful for the  
PT diagnosis and treatment of cancers and immune disorders.  
XX  
Claim 1, Page 1918-1919; 3148pp; English.  
  
This invention describes novel antibodies that immunospecifically bind to  
CC B lymphocyte Stimulator (Blys) polypeptides. Blys is a member of the  
CC tumour necrosis factor (TNF) super family and induces B cell  
CC proliferation and differentiation. The antibodies of the invention have  
CC cytostatic, immunosuppressive, immunostimulant, immunomodulatory,  
CC antirheumatic and antiAIDS activity and can be used in vaccines to  
CC inhibit the expression and activity of Blys. The antibodies bind to Blys  
CC and so may be used to detect and quantitate the presence of Blys in  
CC biological samples and may be used in this way to diagnose disease  
CC associated with aberrant expression of Blys. They may also be

CC administered to treat diseases associated with aberrant BlyS expression  
CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention

XX Sequence 259 AA;

XX Query Match 7.7%; Score 20; DB 5; Length 259;

XX Best Local Similarity 100.0%; Pred. No. 2.9e-09;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SGAEVKKKPGSSVVKVSCKASG 26

Db 7 SGAEVKKKPGSSVVKVSCKASG 26

#### RESULT 8

ID ABP45486 standard; protein; 259 AA.

AC ABP45486;

DT 19-AUG-2002 (first entry)

DE Human BlyS binding scFv SEQ ID 1497.

XX BlyS; B lymphocyte stimulator; TNF superfamily; human; cytosolic;  
XX tumour necrosis factor; B cell proliferation; B cell differentiation;  
XX immunosuppressive; immunostimulant; immunomodulatory; antirheumatic;  
XX antiAIDS; vaccine; cancer; immune; autoimmune disorder; immunodeficiency;  
XX systemic lupus erythematosus; rheumatoid arthritis; CVID; AIDS;  
XX common variable immunodeficiency; acquired immunodeficiency syndrome.

XX Homo sapiens.

XX WO200202641-A1.

XX 10-JAN-2002.

XX 15-JUN-2001; 2001WO-US019110.

XX 16-JUN-2000; 2000US-0212210P.

XX 17-OCT-2000; 2000US-0240816P.

XX 16-MAR-2001; 2001US-0276248P.

XX 21-MAR-2001; 2001US-0277379P.

XX 25-MAY-2001; 2001US-0293499P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (CMB-) CAMBRIDGE ANTIBODY TECHNOLOGY.

XX Ruben SM, Barash SC, Choi GH, Vaughan T, Hilbert D;

XX WPI; 2002-114799/15.

XX Antibodies against B lymphocyte stimulating polypeptides, useful for the

XX diagnosis and treatment of cancers and immune disorders.

XX Claim 1; Page 2189-2190; 3148pp; English.

XX This invention describes novel antibodies that immunospecifically bind to  
XX B lymphocyte stimulator (BlyS) polypeptides. BlyS is a member of the  
XX tumour necrosis factor (TNF) super family and induces B cell  
XX proliferation and differentiation. The antibodies of the invention have  
XX cytosolic, immunosuppressive, immunostimulant, immunomodulatory,  
XX antirheumatic and antiAIDS activity and can be used in vaccines to  
XX inhibit the expression and activity of BlyS. The antibodies bind to BlyS  
XX and so may be used to detect and quantitate the presence of BlyS in  
XX biological samples and may be used in this way to diagnose disease  
XX associated with aberrant expression of BlyS. They may also be  
XX administered to treat diseases associated with aberrant BlyS expression

CC and activity such as cancer, immune, and autoimmune disorders and  
CC diseases, e.g. systemic lupus erythematosus, rheumatoid arthritis,  
CC immunodeficiency (e.g. common variable immunodeficiency (CVID) and  
CC acquired immunodeficiency syndrome (AIDS)). ABP43990-ABP47228 represent  
CC the antibodies and fragments of the antibodies described in the method of  
CC the invention

XX Sequence 259 AA;

XX Query Match 7.7%; Score 20; DB 5; Length 259;

XX Best Local Similarity 100.0%; Pred. No. 2.9e-09;

XX Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SGAEVKKKPGSSVVKVSCKASG 26

Db 7 SGAEVKKKPGSSVVKVSCKASG 26

#### RESULT 9

ID ADG96086 standard; protein; 259 AA.

AC ADG96086;

DT 11-MAR-2004 (first entry)

DE Single chain antibody that immunospecifically binds BlyS Seqid 1270.

XX antibody; B lymphocyte stimulator; BlyS; tumour necrosis factor;  
XX B cell proliferation; differentiation; scFv; myasthenia gravis;  
XX multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
XX carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
XX antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

XX WO2003055979-A2.

XX 10-JUL-2003.

XX 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

XX 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

XX Novel antibody that immunospecifically binds to a B lymphocyte stimulator

XX (BlyS), useful for detecting and treating diseases or disorders e.g.

XX rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1270; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
XX to B lymphocyte stimulator (BlyS). The BlyS gene has been mapped to  
XX chromosome 13q34 and encodes a protein that is a member of the tumour  
XX necrosis factor superfamily and induces both in vivo and in vitro B cell  
XX proliferation and differentiation. Specifically, it refers to single  
XX chain antibody molecules (scFvs) derived, preferably, from the variable  
XX heavy CDR3 region that immunospecifically bind to a polypeptide, or  
XX fragment thereof, of either human, murine, rat or monkey BlyS. The  
XX present invention refers to the use of such antibodies in various methods  
XX for the detection, diagnosis and prognosis of diseases related to the  
XX aberrant expression or inappropriate function of BlyS or its receptor. As  
XX such, these compositions are useful for identifying immune disorders  
XX including myasthenia gravis and multiple sclerosis, inflammatory  
XX disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
XX as AIDS and proliferative disorders including leukaemia, carcinoma and  
XX lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 259 AA;  
  
Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26  
  
RESULT 10  
ADG96091  
ID ADG96091 standard; protein; 259 AA.  
XX  
AC ADG96091;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Single chain antibody that immunospecifically binds Blys SeqID 1275.  
XX  
KM antibody: B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukemia;  
KM carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.  
XX  
OS Unidentified.  
XX  
PN WO2003055979-A2.  
XX  
PD 10-JUL-2003.  
XX  
PF 14-NOV-2002; 2002WO-US036496.  
XX  
PR 16-NOV-2001; 2001US-0331469P.  
PR 19-DEC-2001; 2001US-0340817P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;  
XX  
DR WPI; 2003-505530/47.  
XX  
PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (Blys), useful for detecting and treating diseases or disorders e.g.  
PT rheumatoid arthritis, asthma and leukemia.  
XX  
PS Example 1; SEQ ID NO 1275; 394pp; English.  
XX  
CC This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scfvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.  
XX  
SQ Sequence 259 AA;  
  
Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 7 SGAEVKKPGSSVVKVSCKASG 26  
Db 7 SGAEVKKPGSSVVKVSCKASG 26  
  
RESULT 11  
ADG96313  
ID ADG96313 standard; protein; 259 AA.  
XX  
AC ADG96313;  
XX  
DT 11-MAR-2004 (first entry)  
XX  
DE Single chain antibody that immunospecifically binds Blys SeqID 1497.  
XX  
KM antibody: B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukemia;  
KM carcinoma; lymphoma; antirheumatic; antiallergic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.  
XX  
OS Unidentified.  
XX  
PN WO2003055979-A2.  
XX  
PD 10-JUL-2003.  
XX  
PF 14-NOV-2002; 2002WO-US036496.  
XX  
PR 16-NOV-2001; 2001US-0331469P.  
PR 19-DEC-2001; 2001US-0340817P.  
XX  
PA (HUMA-) HUMAN GENOME SCI INC.  
XX  
PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;  
XX  
DR WPI; 2003-505530/47.  
XX  
PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
PT (Blys), useful for detecting and treating diseases or disorders e.g.  
PT rheumatoid arthritis, asthma and leukemia.  
XX  
PS Example 1; SEQ ID NO 1497; 394pp; English.  
XX  
CC This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scfvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

RESULT 12

ADG96063  
ID ADG96063 standard; protein; 259 AA.

XX ADG96063;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SegID 1247.

XX antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KM carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

OS Unidentified.

PN WO2003055979-A2.

XX 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

DR Novel antibody that immunospecifically binds to a B lymphocyte stimulator

XX (Blys), useful for detecting and treating diseases or disorders e.g.

PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1247; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scfvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritic, neuroprotective,  
CC antiinflammatory, antiasthmatic, antiallergic and cytostatic. This  
CC polypeptide sequence is a single chain antibody that binds Blys of the  
CC invention. NOTE: The sequence data for this patent did not form part of  
CC the printed specification, but was obtained in electronic format  
CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYVSCKASG 26  
Db 7 SGAEVKKPGSSVYVSCKASG 26

RESULT 13

ADG95859  
ID ADG95859 standard; protein; 259 AA.

XX ADG95859;

DT 11-MAR-2004 (first entry)

XX Single chain antibody that immunospecifically binds Blys SegID 1043.

XX antibody; B lymphocyte stimulator; Blys; tumour necrosis factor;  
KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
KM carcinoma; lymphoma; antirheumatic; antiarthritic; neuroprotective;  
KM antiinflammatory; antiasthmatic; antiallergic; cytostatic.

XX Unidentified.

OS Unidentified.

PN WO2003055979-A2.

XX 10-JUL-2003.

PD 14-NOV-2002; 2002WO-US036496.

XX 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

XX (HUMA-) HUMAN GENOME SCI INC.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

XX WPI; 2003-505530/47.

DR Novel antibody that immunospecifically binds to a B lymphocyte stimulator

XX (Blys), useful for detecting and treating diseases or disorders e.g.

PT rheumatoid arthritis, asthma and leukemia.

XX Example 1; SEQ ID NO 1043; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
CC chromosome 13q34 and encodes a protein that is a member of the tumour  
CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
CC proliferation and differentiation. Specifically, it refers to single  
CC chain antibody molecules (scfvs) derived, preferably, from the variable  
CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
CC fragment thereof, of either human, murine, rat or monkey Blys. The  
CC present invention refers to the use of such antibodies in various methods  
CC for the detection, diagnosis and prognosis of diseases related to the  
CC aberrant expression or inappropriate function of Blys or its receptor. As  
CC such, these compositions are useful for identifying immune disorders  
CC including myasthenia gravis and multiple sclerosis, inflammatory  
CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritis, neuroprotective,  
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at fep.wipo.int/pub/published pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKSCKASG 26  
 Db 7 SGAEVKKPGSSVYKSCKASG 26

RESULT 14

ADG96079 standard; protein; 259 AA.

AC ADG96079;

DT 11-MAR-2004 (first entry)

DE Single chain antibody that immunospecifically binds Blys SegID 1263.

XX antibody; B lymphocyte stimulator; Blys; tumor necrosis factor;  
 KM B cell proliferation; differentiation; scfv; myasthenia gravis;  
 KM multiple sclerosis; asthma; rheumatoid arthritis; AIDS; leukaemia;  
 KM carcinoma; lymphoma; antirheumatic; antiarthritis; neuroprotective;  
 KM antiinflammatory; antiaesthetic; antiallergic; cytostatic.

OS Unidentified.

PN WO2003055979-A2.

PD 10-JUL-2003.

PF 14-NOV-2002; 2002WO-US036496.

PR 16-NOV-2001; 2001US-0331469P.

PR 19-DEC-2001; 2001US-0340817P.

PA (HUMA-) HUMAN GENOME SCI INC.

PI Ruben SM, Barash SC, Choi GH, Vaughan TJ, Hilbert D;

DR WPI; 2003-505530/47.

PT Novel antibody that immunospecifically binds to a B lymphocyte stimulator  
 PT (Blys), useful for detecting and treating diseases or disorders e.g.  
 PT rheumatoid arthritis, asthma and leukemia.

PS Example 1; SEQ ID NO 1263; 394pp; English.

XX This invention relates to novel antibodies that immunospecifically bind  
 CC to B lymphocyte stimulator (Blys). The Blys gene has been mapped to  
 CC chromosome 13q34 and encodes a protein that is a member of the tumor  
 CC necrosis factor superfamily and induces both in vivo and in vitro B cell  
 CC proliferation and differentiation. Specifically, it refers to single  
 CC chain antibody molecules (scfvs) derived, preferably, from the variable  
 CC heavy CDR3 region that immunospecifically bind to a polypeptide, or  
 CC fragment thereof, of either human, murine, rat or monkey Blys. The  
 CC present invention refers to the use of such antibodies in various methods  
 CC for the detection, diagnosis and prognosis of diseases related to the  
 CC aberrant expression or inappropriate function of Blys or its receptor. As  
 CC such, these compositions are useful for identifying immune disorders  
 CC including myasthenia gravis and multiple sclerosis, inflammatory  
 CC disorders e.g. asthma and rheumatoid arthritis, infectious diseases such  
 CC as AIDS and proliferative disorders including leukaemia, carcinoma and  
 CC lymphoma. Accordingly, they can be described as exhibiting various

CC activities such as antirheumatic, antiarthritis, neuroprotective,  
 CC antiinflammatory, antiaesthetic, antiallergic and cytostatic. This  
 CC polypeptide sequence is a single chain antibody that binds Blys of the  
 CC invention. NOTE: The sequence data for this patent did not form part of  
 CC the printed specification, but was obtained in electronic format  
 CC directly from WIPO at fep.wipo.int/pub/published pct\_sequences.

XX Sequence 259 AA;

Query Match 7.7%; Score 20; DB 7; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 2.9e-09;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKSCKASG 26  
 Db 7 SGAEVKKPGSSVYKSCKASG 26

RESULT 15

AAU72862 standard; protein; 259 AA.

AC AAU72862;

DT 26-FEB-2002 (first entry)

DE P4-3 single chain Fv.

XX Human; NKGD2; NKGD2 receptor complex; cancer; infectious disease; tumour;  
 KM autoimmune disease; head; neck; stomach; oesophagus; colon; liver; ovary;  
 KM intrahepatic bile duct; pancreas; lung; larynx; breast; uterus; cervix;  
 KM prostate; kidney; testis; thyroid; bladder; brain; melanoma; myeloma; Fv;  
 KM sarcoma; leukaemia; lymphoma; virus; bacterium; fungus; protozoan; DAPI10;  
 KM helminth; cytostatic; antimicrobial; immunomodulatory; 11B2D10; 6H7E7;  
 KM 8G7C10; 6E5A7; 11B2D10x4-7; 8G7C10x4-7; 6E5A7x4-7; P4-2; P4-3; P4-14;  
 KM P4-15; P5-2; P5-3; P5-9; P5-10; P5-11; P5-23; 3B10xP4-3; 3B10xP4-14;  
 KM P53 tetramerisation domain; 3B10xP5-2; 3B10xP5-23.

OS Homo sapiens.

PN WO200171005-A2.

PD 27-SEP-2001.

PF 26-MAR-2001; 2001WO-EP003414.

PR 24-MAR-2000; 2000EP-00106467.

PA (KUFER) KUFER P.

PI Kufer P, Riethmuller G, Luttrebuee R, Borschert K, Kischel R;

PI Mayer M, Hofmeister R;

DR WPI; 2002-055119/07.

DR N-PSDB; AAS97136.

PT Multifunctional polypeptides comprising binding sites that specifically  
 PT recognise extracellular groups of the NKGD2 receptor complex and domains  
 PT which function as receptors or ligands, useful for treating cancers and  
 PT infectious diseases.

PS Example 7; Fig 16; 144pp; English.

XX The invention relates to a multifunctional polypeptide comprising a  
 CC domain with a binding site that specifically recognises an extracellular  
 CC group of the NKGD2 receptor complex and a second domain which functions  
 CC as a receptor or ligand. The polypeptide and its associated  
 CC polynucleotide are used for the preparation of a pharmaceutical  
 CC composition for the treatment of cancer, infections and/or autoimmune  
 CC conditions. The cancer may be a tumour of the head and neck, stomach,  
 CC oesophagus, colon, liver, intrahepatic bile ducts, pancreas, lung,  
 CC larynx, breast, ovary, uterus, cervix, prostate, kidney, testis, thyroid,  
 CC bladder or brain, or a melanoma, myeloma, sarcoma, leukaemia or lymphoma.

CC The infectious diseases can be caused by viruses, bacteria, fungi,  
CC protozoa or helminths. The autoimmune diseases include multiple  
CC sclerosis, Grave's disease, ankylosing spondylitis, acute anterior  
CC uveitis, Goodpasture's syndrome, myasthenia gravis, insulin-dependent  
CC diabetes mellitus, rheumatoid arthritis, pemphigus vulgaris and  
CC autoimmune hepatitis. Sequences AAU72820-AAU72875 represent the NKGD  
CC receptor and the polypeptides of the invention  
XX

SQ Sequence 259 AA;

Query Match 6.9%; Score 18; DB 5; Length 259;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 128 SGGGSGGGSGGGSGSEL 145

Db 119 SGGGSGGGSGGGSGSEL 136

Search completed: September 26, 2005, 13:11:54  
Job time : 162 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:06:17 ; Search time 41 Seconds  
(without alignments)  
471.564 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259

Sequence: 1 QVQLLSGAEVKKRSGSVK.....RYFGQGTRLRIKLVRGSG 259

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 513545 seqs, 74649064 residues

Word size : 0

Total number of hits satisfying chosen parameters: 384

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/1/iaa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/iaa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/iaa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/iaa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/iaa/PTUG\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/iaa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	16	6.2	259	4 US-09-419-788-29	Sequence 29, Appl
2	10	3.9	259	4 US-09-419-788-115	Sequence 115, App
3	7	2.7	259	3 US-08-469-318-141	Sequence 141, App
4	7	2.7	259	3 US-08-468-609A-141	Sequence 141, App
5	7	2.7	259	3 US-09-216-295-5	Sequence 5, Appl
6	7	2.7	259	3 US-08-446-872A-141	Sequence 141, App
7	7	2.7	259	4 US-08-762-227A-141	Sequence 141, App
8	7	2.7	259	4 US-09-632-570-5	Sequence 5, Appl
9	7	2.7	259	4 US-09-632-575-35	Sequence 35, Appl
10	7	2.7	259	5 PCT-US95-01185-141	Sequence 141, App
11	6	2.3	259	4 US-09-079-723-228	Sequence 228, App
12	6	2.3	259	4 US-09-079-723-236	Sequence 236, App
13	6	2.3	259	4 US-09-079-723-243	Sequence 243, App
14	6	2.3	259	4 US-09-079-723-248	Sequence 248, App
15	6	2.3	259	4 US-09-419-788-114	Sequence 114, App
16	5	1.9	259	1 US-07-795-8598-4	Sequence 4, Appl
17	5	1.9	259	1 US-08-457-616-4	Sequence 4, Appl
18	5	1.9	259	2 US-08-408-095-29	Sequence 29, Appl
19	5	1.9	259	3 US-08-469-318-133	Sequence 133, App
20	5	1.9	259	3 US-08-469-318-155	Sequence 155, App
21	5	1.9	259	3 US-08-468-609A-133	Sequence 133, App
22	5	1.9	259	3 US-08-468-609A-155	Sequence 155, App
23	5	1.9	259	3 US-09-082-090-2	Sequence 2, Appl
24	5	1.9	259	3 US-08-446-872A-133	Sequence 133, App
25	5	1.9	259	3 US-08-446-872A-155	Sequence 155, App
26	5	1.9	259	3 US-09-134-001C-3619	Sequence 3619, App
27	5	1.9	259	3 US-09-253-316-27	Sequence 27, Appl

28	5	1.9	259	4 US-08-762-227A-133	Sequence 133, App
29	5	1.9	259	4 US-08-762-227A-155	Sequence 155, App
30	5	1.9	259	4 US-09-252-991A-25689	Sequence 25689, A
31	5	1.9	259	4 US-09-252-991A-28573	Sequence 28573, A
32	5	1.9	259	4 US-09-252-991A-30105	Sequence 30105, A
33	5	1.9	259	4 US-09-742-693-33	Sequence 33, Appl
34	5	1.9	259	4 US-09-489-039A-9829	Sequence 9829, App
35	5	1.9	259	4 US-09-445-774-27	Sequence 27, Appl
36	5	1.9	259	4 US-09-976-594-229	Sequence 229, App
37	5	1.9	259	4 US-09-071-252-9	Sequence 9, Appl
38	5	1.9	259	4 US-09-602-787A-476	Sequence 476, App
39	5	1.9	259	4 US-09-270-767-33541	Sequence 33541, A
40	5	1.9	259	4 US-09-919-039-109	Sequence 109, App
41	5	1.9	259	4 US-09-248-796A-15939	Sequence 15939, A
42	5	1.9	259	4 US-09-248-796A-18507	Sequence 18507, A
43	5	1.9	259	4 US-09-248-796A-20465	Sequence 20465, A
44	5	1.9	259	4 US-09-949-016-92673	Sequence 9267, App
45	5	1.9	259	4 US-09-949-016-10093	Sequence 10093, A

#### ALIGNMENTS

```
RESULT 1
US-09-419-788-29
; Sequence 29, Application US/09419788
; Patent No. 6825325
; GENERAL INFORMATION:
; APPLICANT: FISCHER, Rainer
; APPLICANT: SCHILLBERG, Stefan
; APPLICANT: NAHRING, Jorg
; APPLICANT: SACK, Markus
; APPLICANT: MONECKE, Michael
; APPLICANT: LIAO, Yu-Cai
; APPLICANT: SPIEGEL, Holger
; APPLICANT: ZIMMERMAN, Sabine
; APPLICANT: EMANS, Neil
; TITLE OF INVENTION: Molecular Pathogenicity Mediated Plant Disease
; FILE REFERENCE: 0147-0189P
; CURRENT APPLICATION NUMBER: US/09/419,788
; CURRENT FILING DATE: 1999-10-18
; EARLIER APPLICATION NUMBER: 98 11 9630.6 EP
; EARLIER FILING DATE: 1998-10-16
; EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA
; EARLIER FILING DATE: 1998-10-16
; NUMBER OF SEQ ID NOS: 163
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 29
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: Description of Artificial Sequence: synthetic, no
; OTHER INFORMATION: natural origin
US-09-419-788-29

Query Match          6.2%; Score 16; DB 4; Length 259;
Best Local Similarity 100.0%; Pred. No. 2.4e-05;
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      128 SGGGSGGGSGGGGS 143
Db       116 SGGGSGGGSGGGGS 131

RESULT 2
US-09-419-788-115
; Sequence 115, Application US/09419788
; Patent No. 6825325
; GENERAL INFORMATION:
; APPLICANT: FISCHER, Rainer
; APPLICANT: SCHILLBERG, Stefan
```

APPLICANT: NAHRING, JORG  
APPLICANT: SACK, Markus  
APPLICANT: MONECKE, Michael  
APPLICANT: LIAO, Yu-Cai  
APPLICANT: SPIEGEL, Holger  
APPLICANT: ZIMMERMAN, Sabine  
APPLICANT: EMANS, Neil  
TITLE OF INVENTION: Molecular Pathogenicity Mediated Plant Disease  
FILE REFERENCE: 0147-0189P  
CURRENT APPLICATION NUMBER: US/09/419,798  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 98 11 9630.6 EP  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 115  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic, no  
OTHER INFORMATION: natural origin  
US-09-419-788-115

Query Match 3.9%; Score 10; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 10; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 LAWYQKPGQ 185  
DB 179 LAWYQKPGQ 188

RESULT 3  
US-08-469-318-141  
Sequence 141, Application US/08469318  
Patent No. 6022535  
GENERAL INFORMATION:  
APPLICANT:  
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
TITLE OF INVENTION: Protein  
NUMBER OF SEQUENCES: 196  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/469,318  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/446,872  
FILING DATE:  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-469-318-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGGG 136  
DB 123 GGGSGGG 129

RESULT 4  
US-08-468-609A-141  
Sequence 141, Application US/08468609A  
Patent No. 6030812  
GENERAL INFORMATION:  
APPLICANT: Abrams, Mark A.  
APPLICANT: Bauer, S. C.  
APPLICANT: Braford-Goldberg, Sarah R.  
APPLICANT: Caparon, Maite H.  
APPLICANT: Easton, Alan M.  
APPLICANT: Klein, Barbara K.  
APPLICANT: McKeown, John P.  
APPLICANT: Olin, Peter O.  
APPLICANT: Paik, Kumman  
APPLICANT: Thomas, John W.  
TITLE OF INVENTION: Fusion Proteins Comprising Multiply Mutated Interleukin-3 (IL-3)  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSER: Dennis A. Bennett, G.D. Searle & Co.,  
ADDRESSER: Corporate Patent Dept.  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/468,609A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/192,325  
FILING DATE: 14-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Dennis A.  
REGISTRATION NUMBER: 34,547  
REFERENCE/DOCKET NUMBER: C-2790/3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6986  
TELEFAX: (314)737-6972  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-468-609A-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGSGGG 136  
DB 123 GGGSGGG 129

RESULT 5  
US-09-216-295-5  
Sequence 5, Application US/09216295  
Patent No. 6268328  
GENERAL INFORMATION:  
APPLICANT: Mitchinson, Colin  
APPLICANT: Wendt, Dan J.  
TITLE OF INVENTION: No. 6268328e1 Variant EGIII-like Cellulase Compositions  
FILE REFERENCE: GC555

CURRENT APPLICATION NUMBER: US/09/216,295  
CURRENT FILING DATE: 1998-12-18  
NUMBER OF SEQ ID NOS: 41  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO: 5  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Aspergillus aculeatus  
US-09-216-295-5

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 151 PATSVS 157  
Db 224 PATSVS 230

## RESULT 6

US-08-446-872A-141  
Sequence 141, Application US/08446872A  
Patent No. 6361977  
GENERAL INFORMATION:  
APPLICANT: Abrams, Mark A.  
APPLICANT: Bauer, S. C.  
APPLICANT: Braford-Goldberg, Sarah R.  
APPLICANT: Caparon, Maite H.  
APPLICANT: Easton, Alan M.  
APPLICANT: Klein, Barbara K.  
APPLICANT: McKearn, John P.  
APPLICANT: Olns, Peter O.  
APPLICANT: Paik, Kumman  
APPLICANT: Thomas, John W.  
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis  
TITLE OF INVENTION: Fusion Protein  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
ADDRESSEE: Corporate Patent Dept.,  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/446,872A  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 424  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/192,325  
FILING DATE: 14-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Dennis A.  
REGISTRATION NUMBER: 34,547  
REFERENCE/DOCKET NUMBER: C-2790/1  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (314)737-6986  
TELEFAX: (314)737-6972  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-446-872A-141

Query Match 2.7%; Score 7; DB 3; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 GGGSGGG 136  
Db 123 GGGSGGG 129

## RESULT 7

US-08-762-227A-141  
Sequence 141, Application US/08762227A  
Patent No. 6436387  
GENERAL INFORMATION:  
APPLICANT: Abrams, Mark A.  
APPLICANT: Bauer, S. C.  
APPLICANT: Braford-Goldberg, Sarah R.  
APPLICANT: Caparon, Maite H.  
APPLICANT: Easton, Alan M.  
APPLICANT: Klein, Barbara K.  
APPLICANT: McKearn, John P.  
APPLICANT: Olns, Peter O.  
APPLICANT: Paik, Kumman  
APPLICANT: Thomas, John W.  
TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis  
TITLE OF INVENTION: Fusion Protein  
NUMBER OF SEQUENCES: 197  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dennis A. Bennett, G.D. Searle & Co.,  
ADDRESSEE: Corporate Patent Dept.,  
STREET: P. O. Box 5110  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60680  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/762,227A  
FILING DATE: 09-Dec-1996  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/192,325  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: US 08/446,872  
FILING DATE: 06-JUN-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: Bennett, Dennis A.  
REGISTRATION NUMBER: 34,547  
REFERENCE/DOCKET NUMBER: C-2790/5  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (708)470-6501  
TELEFAX: (708)470-6881  
INFORMATION FOR SEQ ID NO: 141:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS: <Unknown>  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 141:  
US-08-762-227A-141

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 130 GGGSGGG 136  
Db 123 GGGSGGG 129

RESULT 8  
US-09-632-570-5  
; Sequence 5, Application US/09632570  
; Patent No. 6623949  
; GENERAL INFORMATION:  
; APPLICANT: Gualietti, Peter  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: Phillips, Jay Ian  
; TITLE OF INVENTION: No. 6623949el Variant EGIII-Like Cellulase  
; TITLE OF INVENTION: Compositions  
; FILE REFERENCE: GC631  
; CURRENT APPLICATION NUMBER: US/09/632,570  
; CURRENT FILING DATE: 2000-08-04  
; NUMBER OF SEQ ID NOS: 64  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 5  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Aspergillus aculeatus  
US-09-632-570-5

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PATLSVS 157  
Db 224 PATLSVS 230

RESULT 9  
US-09-632-575-35  
; Sequence 35, Application US/09632575  
; Patent No. 6635465  
; GENERAL INFORMATION:  
; APPLICANT: Gualietti, Peter  
; APPLICANT: Mitchinson, Colin  
; APPLICANT: ROPP, Traci M.  
; TITLE OF INVENTION: Mutant EGIII Cellulase, DNA Encoding  
; TITLE OF INVENTION: Such EGIII Compositions and Methods for Obtaining Same  
; FILE REFERENCE: GC629  
; CURRENT APPLICATION NUMBER: US/09/632,575  
; CURRENT FILING DATE: 2000-08-04  
; NUMBER OF SEQ ID NOS: 54  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 35  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Aspergillus aculeatus  
US-09-632-575-35

Query Match 2.7%; Score 7; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 151 PATLSVS 157  
Db 224 PATLSVS 230

RESULT 10  
PCT-US95-01185-141  
; Sequence 141, Application PC/TUS9501185  
; GENERAL INFORMATION:  
; APPLICANT:  
; TITLE OF INVENTION: Multivariant IL-3 Hematopoiesis Fusion  
; NUMBER OF SEQUENCES: 196  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30 (EPO)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US95/01185  
; FILING DATE: 02-FEB-1995  
; CLASSIFICATION:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 08/192325  
; FILING DATE: 14-FEB-1994  
; INFORMATION FOR SEQ ID NO: 141:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US95-01185-141

Query Match 2.7%; Score 7; DB 5; Length 259;  
Best Local Similarity 100.0%; Pred. No. 4.5e+02;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 130 GGGGGG 136  
Db 123 GGGGGG 129

RESULT 11  
US-09-079-723-228  
; Sequence 228, Application US/09079723  
; Patent No. 6703362  
; GENERAL INFORMATION:  
; APPLICANT: Alvarez, Vernon L.  
; APPLICANT: O'Mahony, Daniel J.  
; APPLICANT: Lambkin, Imelda J.  
; APPLICANT: Singleton, Judith  
; APPLICANT: Paterson, Catherine A.  
; APPLICANT: Cagney, Gerard M.  
; APPLICANT: Bellinka, Benjamin A.  
; APPLICANT: Carter, John M.  
; TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-  
; TITLE OF INVENTION: INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS  
; NUMBER OF SEQUENCES: 265  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds LLP  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/079,723  
; FILING DATE:  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Miarock, S. Leslie  
; REGISTRATION NUMBER: 18, 872  
; REFERENCE/DOCKET NUMBER: 1101-219  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212-790-9090  
; TELEFAX: 212-869-9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 228:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 259 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: unknown

MOLECULE TYPE: peptide  
US-09-079-723-228

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
DB 221 LVPRGS 226

RESULT 12  
US-09-079-723-236

Sequence 236, Application US/09079723

Patent No. 6703362

GENERAL INFORMATION:

APPLICANT: Alvarez, Vernon L.

APPLICANT: O'Mahony, Daniel J.

APPLICANT: Lambkin, Imelda J.

APPLICANT: Singleton, Judith

APPLICANT: Patterson, Catherine A.

APPLICANT: Cagney, Gerard M.

APPLICANT: Belinka, Benjamin A.

APPLICANT: Carter, John M.

TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-

TITLE OF INVENTION: INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/079, 723

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Mierock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-219

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 236:

SEQUENCE CHARACTERISTICS:

LENGTH: 259 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-09-079-723-236

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
DB 221 LVPRGS 226

RESULT 13

US-09-079-723-243

Sequence 243, Application US/09079723

Patent No. 6703362

GENERAL INFORMATION:

APPLICANT: Alvarez, Vernon L.

APPLICANT: O'Mahony, Daniel J.

APPLICANT: Lambkin, Imelda J.

APPLICANT: Singleton, Judith

APPLICANT: Patterson, Catherine A.

APPLICANT: Cagney, Gerard M.

APPLICANT: Belinka, Benjamin A.

APPLICANT: Carter, John M.

TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-

TITLE OF INVENTION: INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York

COUNTRY: USA

ZIP: 10036

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/079, 723

FILING DATE:

CLASSIFICATION:

ATTORNEY/AGENT INFORMATION:

NAME: Mierock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 1101-219

TELECOMMUNICATION INFORMATION:

TELEPHONE: 212-790-9090

TELEFAX: 212-869-9741

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 243:

SEQUENCE CHARACTERISTICS:

LENGTH: 259 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: unknown

MOLECULE TYPE: peptide

US-09-079-723-243

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 253 LVPRGS 258  
|||||  
DB 221 LVPRGS 226

RESULT 14  
US-09-079-723-248

Sequence 248, Application US/09079723

Patent No. 6703362

GENERAL INFORMATION:

APPLICANT: Alvarez, Vernon L.

APPLICANT: O'Mahony, Daniel J.

APPLICANT: Lambkin, Imelda J.

APPLICANT: Singleton, Judith

APPLICANT: Patterson, Catherine A.

APPLICANT: Cagney, Gerard M.

APPLICANT: Belinka, Benjamin A.

APPLICANT: Carter, John M.

TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-

TITLE OF INVENTION: INTESTINAL TRACT (GIT) TRANSPORT RECEPTORS AND RELATED METHODS

NUMBER OF SEQUENCES: 265

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds LLP

STREET: 1155 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/079,723  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mistrock, S. Leslie  
REGISTRATION NUMBER: 18,872  
REFERENCE/DOCKET NUMBER: 1101-219  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-790-9090  
TELEFAX: 212-869-9741  
TELEX: 66141 PENNIE  
INFORMATION FOR SEQ ID NO: 248:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 259 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-09-079-723-248

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 253 LVPRGS 258  
DB 221 LVPRGS 226

RESULT 15  
US-09-419-788-114  
Sequence 114, Application US/09419788  
Patent No. 6825325  
GENERAL INFORMATION:  
APPLICANT: FISCHER, Rainer  
APPLICANT: SCHILLBERG, Stefan  
APPLICANT: NAHRING, Jorg  
APPLICANT: SACK, Markus  
APPLICANT: MONECKE, Michael  
APPLICANT: LIAO, Yu-Cai  
APPLICANT: SPIEGEL, Holger  
APPLICANT: ZIMMERMAN, Sabine  
APPLICANT: EMANS, Neil  
TITLE OF INVENTION: Molecular Pathogenicicid Mediated Plant Disease  
TITLE OF INVENTION: Resistance  
FILE REFERENCE: 0147-0189P  
CURRENT APPLICATION NUMBER: US/09/419,788  
CURRENT FILING DATE: 1999-10-18  
EARLIER APPLICATION NUMBER: 98 11 9630.6 EP  
EARLIER FILING DATE: 1998-10-16  
EARLIER APPLICATION NUMBER: 66/BOM/1998 INDIA  
EARLIER FILING DATE: 1998-10-16  
NUMBER OF SEQ ID NOS: 163  
SOFTWARE: Patentin Ver. 2.1  
SEQ ID NO 114  
LENGTH: 259  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: synthetic, no  
US-09-419-788-114

Query Match 2.3%; Score 6; DB 4; Length 259;  
Best Local Similarity 100.0%; Pred. No. 2.9e+03;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 118 WGGQTL 123  
DB 113 WGGQTL 118

Search completed: September 26, 2005, 13:16:25  
Job time : 42 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 26, 2005, 13:14:58 ; Search time 168 seconds  
(without alignments)  
627.497 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259

Sequence: 1 QVQLRSGAEVKKPKGSSVKY.....RYTFGGTRLIKLIVRSGS 259

Scoring table: OLIGO

Gapop 60.0 , Gapext 60.0

Searched: 1826554 seqs, 407025358 residues

Word size : 0

Total number of hits satisfying chosen parameters: 3391

Minimum DB seq length: 259

Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubppaa/US07\_PUBCOMB.pep:\*

2: /cgn2\_6/ptodata/2/pubppaa/PCT\_NEW\_PUB.pep:\*

3: /cgn2\_6/ptodata/2/pubppaa/US06\_NEW\_PUB.pep:\*

4: /cgn2\_6/ptodata/2/pubppaa/US06\_PUBCOMB.pep:\*

5: /cgn2\_6/ptodata/2/pubppaa/US07\_NEW\_PUB.pep:\*

6: /cgn2\_6/ptodata/2/pubppaa/PCTUS\_PUBCOMB.pep:\*

7: /cgn2\_6/ptodata/2/pubppaa/US08\_NEW\_PUB.pep:\*

8: /cgn2\_6/ptodata/2/pubppaa/US08\_PUBCOMB.pep:\*

9: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*

10: /cgn2\_6/ptodata/2/pubppaa/US09\_PUBCOMB.pep:\*

11: /cgn2\_6/ptodata/2/pubppaa/US09C\_PUBCOMB.pep:\*

12: /cgn2\_6/ptodata/2/pubppaa/US09\_NEW\_PUB.pep:\*

13: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

14: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

15: /cgn2\_6/ptodata/2/pubppaa/US10C\_PUBCOMB.pep:\*

16: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

17: /cgn2\_6/ptodata/2/pubppaa/US10\_PUBCOMB.pep:\*

18: /cgn2\_6/ptodata/2/pubppaa/US10\_NEW\_PUB.pep:\*

19: /cgn2\_6/ptodata/2/pubppaa/US11\_PUBCOMB.pep:\*

20: /cgn2\_6/ptodata/2/pubppaa/US11\_NEW\_PUB.pep:\*

21: /cgn2\_6/ptodata/2/pubppaa/US60\_NEW\_PUB.pep:\*

22: /cgn2\_6/ptodata/2/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	9.7	259	US-10-863-729-18	Sequence 18, Appl
2	21	8.1	259	US-09-880-748-1034	Sequence 1034, Ap
3	21	8.1	259	US-10-293-418-1034	Sequence 1034, Ap
4	20	7.7	259	US-09-880-748-1043	Sequence 1043, Ap
5	20	7.7	259	US-09-880-748-1247	Sequence 1247, Ap
6	20	7.7	259	US-09-880-748-1263	Sequence 1263, Ap
7	20	7.7	259	US-09-880-748-1270	Sequence 1270, Ap
8	20	7.7	259	US-09-880-748-1275	Sequence 1275, Ap
9	20	7.7	259	US-09-880-748-1457	Sequence 1457, Ap
10	20	7.7	259	US-10-293-418-1043	Sequence 1043, Ap
11	20	7.7	259	US-10-293-418-1247	Sequence 1247, Ap

12	20	7.7	259	15	US-10-293-418-1263	Sequence 1263, Ap
13	20	7.7	259	15	US-10-293-418-1270	Sequence 1270, Ap
14	20	7.7	259	15	US-10-293-418-1275	Sequence 1275, Ap
15	20	7.7	259	15	US-10-293-418-1497	Sequence 1497, Ap
16	18	6.9	259	15	US-10-293-656-53	Sequence 53, Appl
17	16	6.2	259	10	US-09-880-748-864	Sequence 864, Appl
18	16	6.2	259	10	US-09-880-748-973	Sequence 973, Appl
19	16	6.2	259	10	US-09-880-748-1025	Sequence 1025, Ap
20	16	6.2	259	10	US-09-880-748-1032	Sequence 1032, Ap
21	16	6.2	259	10	US-09-880-748-1041	Sequence 1041, Ap
22	16	6.2	259	10	US-09-880-748-1262	Sequence 1262, Ap
23	16	6.2	259	10	US-09-880-748-1277	Sequence 1277, Ap
24	16	6.2	259	10	US-09-880-748-1356	Sequence 1356, Ap
25	16	6.2	259	10	US-09-880-748-1452	Sequence 1452, Ap
26	16	6.2	259	10	US-09-880-748-1481	Sequence 1481, Ap
27	16	6.2	259	10	US-09-880-748-1485	Sequence 1485, Ap
28	16	6.2	259	10	US-09-880-748-1552	Sequence 1552, Ap
29	16	6.2	259	10	US-09-880-748-1567	Sequence 1567, Ap
30	16	6.2	259	10	US-09-880-748-1664	Sequence 1664, Ap
31	16	6.2	259	15	US-10-293-418-864	Sequence 864, Appl
32	16	6.2	259	15	US-10-293-418-973	Sequence 973, Appl
33	16	6.2	259	15	US-10-293-418-1025	Sequence 1025, Ap
34	16	6.2	259	15	US-10-293-418-1032	Sequence 1032, Ap
35	16	6.2	259	15	US-10-293-418-1041	Sequence 1041, Ap
36	16	6.2	259	15	US-10-293-418-1262	Sequence 1262, Ap
37	16	6.2	259	15	US-10-293-418-1277	Sequence 1277, Ap
38	16	6.2	259	15	US-10-293-418-1356	Sequence 1356, Ap
39	16	6.2	259	15	US-10-293-418-1452	Sequence 1452, Ap
40	16	6.2	259	15	US-10-293-418-1481	Sequence 1481, Ap
41	16	6.2	259	15	US-10-293-418-1485	Sequence 1485, Ap
42	16	6.2	259	15	US-10-293-418-1552	Sequence 1552, Ap
43	16	6.2	259	15	US-10-293-418-1567	Sequence 1567, Ap
44	16	6.2	259	15	US-10-293-418-1664	Sequence 1664, Ap
45	15	5.8	259	14	US-10-207-655-13	Sequence 13, Appl

#### ALIGNMENTS

RESULT 1

US-10-863-729-18

Sequence 18, Application US/10863729

Publication No. US20050013819A1

GENERAL INFORMATION:

APPLICANT: Kinch, Michael S.

TITLE OF INVENTION: USE OF EPH4 AND MODULATOR OF EPH4 FOR

FILE REFERENCE: 10271-117-999

CURRENT APPLICATION NUMBER: US/10/863,729

CURRENT FILING DATE: 2004-06-07

PRIOR APPLICATION NUMBER: 60/476,909

PRIOR FILING DATE: 2003-06-06

PRIOR APPLICATION NUMBER: 60/503,356

PRIOR FILING DATE: 2003-09-16

NUMBER OF SEQ ID NOS: 32

SOFTWARE: PastSeq for Windows Version 4.0

SEQ ID NO 18

LENGTH: 259

TYPE: PRT

ORGANISM: Homo Sapiens

FEATURE:

OTHER INFORMATION: Mature scFv (EA44) sequence with FLAG & HIS Tags

US-10-863-729-18

Query Match 9.7%; Score 25; DB 17; Length 259;

Best local Similarity 100.0%; Pred. No. 3.7e-13;

Matches 25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 176 LAWYQKPGQAPRLIYGASTRATG 200

DB 164 LAWYQKPGQAPRLIYGASTRATG 188

```
RESULT 2
US-09-880-748-1034
; Sequence 1034, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1034
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1034

Query Match
Best Local Similarity 100.0%; Score 21; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 1e-09;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 6 ESGAEVKKPGSSVVKSCASG 26
Db 6 ESGAEVKKPGSSVVKSCASG 26

RESULT 3
US-10-293-418-1034
; Sequence 1034, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1034
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1034

Query Match
Best Local Similarity 100.0%; Score 21; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 1e-09;

Qy 7 SGAEVKKPGSSVVKSCASG 26
Db 7 SGAEVKKPGSSVVKSCASG 26

RESULT 4
US-09-880-748-1043
; Sequence 1043, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1043
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1043

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 7 SGAEVKKPGSSVVKSCASG 26
Db 7 SGAEVKKPGSSVVKSCASG 26

RESULT 5
US-09-880-748-1247
; Sequence 1247, Application US/09880748
; Publication No. US20030059937A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: PF523
; CURRENT APPLICATION NUMBER: US/09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; NUMBER OF SEQ ID NOS: 3239
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1247
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-880-748-1247

Query Match
Best Local Similarity 100.0%; Score 20; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.7%;
```

Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

## RESULT 6

US-09-880-748-1263  
; Sequence 1263, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1263  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1263

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

## RESULT 7

US-09-880-748-1270  
; Sequence 1270, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1270  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1270

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

## RESULT 8

US-09-880-748-1275  
; Sequence 1275, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1275  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1275

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVYKVCCKASG 26  
Db 7 SGAEVKKPGSSVYKVCCKASG 26

## RESULT 9

US-09-880-748-1497  
; Sequence 1497, Application US/09880748  
; Publication No. US20030059937A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523  
; CURRENT APPLICATION NUMBER: US/09/880,748  
; CURRENT FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-15  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; NUMBER OF SEQ ID NOS: 3239  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 1497  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-09-880-748-1497

Query Match 7.7%; Score 20; DB 10; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 10

US-10-293-418-1043  
; Sequence 1043, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1043  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1043

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 11

US-10-293-418-1247  
; Sequence 1247, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248

; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1247  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1247

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 12

US-10-293-418-1263  
; Sequence 1263, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27  
; PRIOR APPLICATION NUMBER: 60/331,469  
; PRIOR FILING DATE: 2001-11-16  
; PRIOR APPLICATION NUMBER: 60/340,817  
; PRIOR FILING DATE: 2001-12-19  
; PRIOR APPLICATION NUMBER: 09/880,748  
; PRIOR FILING DATE: 2001-06-15  
; PRIOR APPLICATION NUMBER: 60/293,499  
; PRIOR FILING DATE: 2001-05-25  
; PRIOR APPLICATION NUMBER: 60/277,379  
; PRIOR FILING DATE: 2001-03-21  
; PRIOR APPLICATION NUMBER: 60/276,248  
; PRIOR FILING DATE: 2001-03-16  
; PRIOR APPLICATION NUMBER: 60/240,816  
; PRIOR FILING DATE: 2000-10-17  
; PRIOR APPLICATION NUMBER: 60/212,210  
; PRIOR FILING DATE: 2000-06-16  
; NUMBER OF SEQ ID NOS: 3247  
; SEQ ID NO 1263  
; LENGTH: 259  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-293-418-1263

Query Match 7.7%; Score 20; DB 15; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.5e-09;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 7 SGAEVKKPGSSVSVKSCASG 26  
|||||  
DB 7 SGAEVKKPGSSVSVKSCASG 26

RESULT 13

US-10-293-418-1270  
; Sequence 1270, Application US/10293418  
; Publication No. US20030223996A1  
; GENERAL INFORMATION:  
; APPLICANT: Ruben et al.  
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys  
; FILE REFERENCE: PF523P2  
; CURRENT APPLICATION NUMBER: US/10/293,418  
; PRIOR FILING DATE: 2002-11-27

```

; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1270
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1270
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
       |||
Db      7 SGAEVKKPGSSVVKSCKASG 26
```

```

RESULT 14
US-10-293-418-1275
; Sequence 1275, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: P5523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1275
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1275
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
       |||
Db      7 SGAEVKKPGSSVVKSCKASG 26
```

```

RESULT 15
US-10-293-418-1497
; Sequence 1497, Application US/10293418
; Publication No. US20030223996A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Antibodies that Immunospecifically Bind Blys
; FILE REFERENCE: P5523P2
; CURRENT APPLICATION NUMBER: US/10/293,418
; PRIOR FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/331,469
; PRIOR FILING DATE: 2001-11-16
; PRIOR APPLICATION NUMBER: 60/340,817
; PRIOR FILING DATE: 2001-12-19
; PRIOR APPLICATION NUMBER: 09/880,748
; PRIOR FILING DATE: 2001-06-15
; PRIOR APPLICATION NUMBER: 60/293,499
; PRIOR FILING DATE: 2001-05-25
; PRIOR APPLICATION NUMBER: 60/277,379
; PRIOR FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: 60/276,248
; PRIOR FILING DATE: 2001-03-16
; PRIOR APPLICATION NUMBER: 60/240,816
; PRIOR FILING DATE: 2000-10-17
; PRIOR APPLICATION NUMBER: 60/212,210
; PRIOR FILING DATE: 2000-06-16
; NUMBER OF SEQ ID NOS: 3247
; SEQ ID NO 1497
; LENGTH: 259
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-293-418-1497
```

```
Query Match      7.7%; Score 20; DB 15; Length 259;
Best Local Similarity 100.0%; Pred. No. 7.5e-09;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY      7 SGAEVKKPGSSVVKSCKASG 26
       |||
Db      7 SGAEVKKPGSSVVKSCKASG 26
```

```
Search completed: September 26, 2005, 13:28:36
Job time : 168 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 26, 2005, 13:05:27 ; Search time 40 Seconds  
(without alignments)  
623.004 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259  
Sequence: 1 QVQLPSGAEVKKRSGSVK.....RYTFGGTRLEIKLVPRGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 283416 seqs, 96216763 residues

Word size : 0

Total number of hits satisfying chosen parameters: 603

Minimum DB seq length: 259  
Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.3	259	2 T15126	hypothetical prote
2	6	2.3	259	2 C84165	hypothetical prote
3	6	2.3	259	2 F86475	hypothetical prote
4	5	1.9	259	1 C47069	acetolactate decar
5	5	1.9	259	1 IOHUI	insulin-like growt
6	5	1.9	259	1 S46286	RNA-binding protei
7	5	1.9	259	2 D49343	cis-1,2-dihydro-1,
8	5	1.9	259	2 F83479	probable short-cha
9	5	1.9	259	2 B55217	cis-1,2-dihydro-1,
10	5	1.9	259	2 B83838	oxidoreductase BHI
11	5	1.9	259	2 C75410	methionine aminope
12	5	1.9	259	2 T28147	class II histocomp
13	5	1.9	259	2 T48119	P-glycoprotein - C
14	5	1.9	259	2 AC0358	probable pili chap
15	5	1.9	259	2 D75372	probable nitrogen
16	5	1.9	259	2 AF1847	hypothetical prote
17	5	1.9	259	2 G97397	ubiquitinone/menqui
18	5	1.9	259	2 T37185	probable gntR-fami
19	5	1.9	259	2 AC3100	ABC transporter, m
20	5	1.9	259	2 G98186	hypothetical prote
21	5	1.9	259	2 C95939	probable spermidin
22	5	1.9	259	2 B64514	hypothetical prote
23	5	1.9	259	2 AF0999	conserved hypotet
24	5	1.9	259	2 A55274	3-oxo-5alpha-stero
25	5	1.9	259	2 S65744	steroid 5alpha-red
26	5	1.9	259	2 F81877	hypothetical prote
27	5	1.9	259	2 B84886	probable beta-expa
28	5	1.9	259	2 T50657	beta-expansin [imp
29	5	1.9	259	2 T13260	hypothetical prote

30	5	1.9	259	2 D86685	prophage p11 prote
31	5	1.9	259	2 C86797	prophage p13 prote
32	5	1.9	259	2 C86757	prophage p12 prote
33	5	1.9	259	2 T23782	hypothetical prote
34	5	1.9	259	2 G96673	hypothetical prote
35	5	1.9	259	2 E87537	hypothetical prote
36	5	1.9	259	2 T51745	RNA helicase RH16
37	5	1.9	259	2 D84343	TRK potassium upla
38	5	1.9	259	2 T17889	glycine-rich prote
39	5	1.9	259	2 E69811	hypothetical prote
40	5	1.9	259	2 JE0326	peroxin Pex1lp iso
41	5	1.9	259	2 C97664	hypothetical prote
42	4	1.5	259	1 DECSF	sorbitol-6-phospha
43	4	1.5	259	1 YXBYT	thymidylate syntha
44	4	1.5	259	1 XYBPT2	site-specific DNA-
45	4	1.5	259	1 XYBPT4	site-specific DNA-

#### ALIGNMENTS

##### RESULT 1

T15126  
hypothetical protein T20B6.3 - *Caenorhabditis elegans*

C/Species: *Caenorhabditis elegans*

C/Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C/Accession: T15126

R/Beck, C.; Wamsley, P.

submitted to the EMBL Data Library, April 1997

A/Description: The sequence of C. elegans coemid T20B6.

A/Reference number: Z18297

A/Accession: T15126

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-259 <BEC>

A/Cross-references: UNIPROT:O02049; EMBL:AF000193; NID:g1946986; PID:g1946988; PIDN:AA85;

A/Experimental source: strain Bristol N2; clone T20B6

C/Genetics:

A/Gene: CESP:T20B6.3

A/Map position: 3

A/Inserts: 9/2; 231/1

C/Superfamily: Phaseolus glycine-rich protein 1.0

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 GSGGGG 136  
Db 80 GSGGGG 85

##### RESULT 2

C84165  
hypothetical protein Vng0049h [imported] - *Halobacterium* sp.

C/Species: *Halobacterium* sp. NRC-1

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: C84165

R/Ng, W.V.; Kennedy, S.P.; Mahairas, G.G.; Berguist, B.; Pan, M.; Shukla, H.D.; Lasky, S.

; Leitbauer, B.; Keller, K.; Cruz, R.; Dawson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablor

Jung, K.H.; Alam, M.; Freltas, T.

Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000

A/Title: Genome sequence of *Halobacterium* species NRC-1.

A/Reference number: A84160; WUID:20504483; PMID:11016950

A/Accession: C84165

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-259 <STO>

A/Cross-references: UNIPROT:Q9HSM3; GB:AE004437; NID:g105796699; PIDN:AA818687.1; GSPDB:GT

C/Genetics:

A/Gene: VNG0049H

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GSSVKV 20  
|||  
Db 134 GSSVKV 139

RESULT 3  
F86475  
hypothetical protein F12A4.2 - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: F86475  
R/Theologas, A.; Becker, J.R.; Palm, C.J.; Federapfel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; anson, N.F.; Hughes, B.; Hitzar, L.  
Nature 408, 816-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lutro, J.S.; Maitl, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salzberg, S.L.; Schwartz, J.R.; Shim, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; PMID:2106719; PMID:11130712  
A/Accession: F86475  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <STO>  
A/Cross-references: UNIPROT:Q9C6P8; GB:A0051172; NID:g10092190; PIDN:AA612608.1; GSTDB:C  
C/Genetic:  
A/Map position: 1

Query Match 2.3%; Score 6; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 8.8e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 206 SSGSGG 211  
|||||  
Db 107 SSGSGG 112

## RESULT 4

C47069  
acetylactate decarboxylase (EC 4.1.1.5) - Klebsiella terrigena

C/Species: Klebsiella terrigena  
C/Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
C/Accession: C47069  
R/Blomqvist, K.; Nikoila, M.; Lehtovaara, P.; Suikko, M.L.; Airaksinen, U.; Straby, K.B.; J. Bacteriol. 175, 1392-1404, 1993  
A/Title: Characterization of the genes of the 2,3-butanediol operons from Klebsiella terrigena  
A/Reference number: A47069; PMID:8444801  
A/Accession: C47069  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <BIO>  
A/Cross-references: UNIPROT:Q04518; GB:I04507; NID:g149170; PIDN:AAA25054.1; PID:g149171  
A/Experimental source: VTT-E-74023  
A/Note: sequence extracted from NCBI backbone (NCBI:126766, NCBI:P.126767)  
C/Superfamily: acetylactate decarboxylase  
C/Keyword: carbon-carbon lyase; carboxy-lyase

Query Match 1.9%; Score 5; DB 1; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 98 GYEG 102  
|||||  
Db 41 GYEG 45

## RESULT 5

IOH1  
Insulin-like growth factor-binding protein 1 precursor [validated] - human  
N/Alternate names: IGF-binding protein; IGF-BP25; insulin-like growth factor 25k binding  
C/Species: Homo sapiens (man)  
C/Date: 30-Sep-1992 #sequence\_revision 30-Sep-1992 #text\_change 09-Jul-2004  
C/Accession: A31867; S01159; A40155; A40925; A42176; S43207; S03559; S01333; A27694; S014  
R/Brinkman, A.; Groffen, C.A.H.; Kortleve, D.J.; Drop, S.L.S.  
Biochem. Biophys. Res. Commun. 157, 898-907, 1988  
A/Title: Organization of the gene encoding the insulin-like growth factor binding protein  
A/Reference number: A31867; PMID:89087480; PMID:2849945  
A/Accession: A31867  
A/Molecule type: DNA  
A/Residues: 1-259 <BR>  
A/Cross-references: UNIPROT:P08833; GB:M23592; GB:M23593; GB:M23594; GB:M23595; NID:g1844  
R/Brinkman, A.; Groffen, C.; Kortleve, D.J.; Geurts van Kessel, A.; Drop, S.L.S.  
EMBO J. 7, 2417-2423, 1988  
A/Title: Isolation and characterization of a cDNA encoding the low molecular weight insu  
A/Reference number: S01159; PMID:89052654; PMID:2461294  
A/Accession: S01159  
A/Molecule type: mRNA  
A/Residues: 1-252, 'W', 254-259 <BR2>  
A/Cross-references: EMBL:Y00856; NID:g32608; PIDN:CAA68770.1; PID:g32609  
A/Note: 169-Thr was also found  
R/Cubbage, M.L.; Suwanichkul, A.; Powell, D.R.  
Mol. Endocrinol. 3, 846-851, 1989  
A/Title: Structure of the human chromosome1 gene for the 25 kilodalton insulin-like growth  
A/Reference number: A40155; PMID:89330502; PMID:2474129  
A/Accession: A40155  
A/Molecule type: DNA  
A/Residues: 1-259 <CUB>  
A/Cross-references: GB:M59316; NID:g184809; PIDN:AAA52793.1; PID:g184810  
R/Lee, Y.L.; Hinz, R.B.; James, P.M.; Lee, P.D.K.; Shively, J.E.; Powell, D.R.  
Mol. Endocrinol. 2, 404-411, 1988  
A/Title: Insulin-like growth factor (IGF) binding protein complementary deoxyribonucleic  
A/Reference number: A40925; PMID:89334540; PMID:2458522  
A/Accession: A40925  
A/Molecule type: mRNA  
A/Residues: 1-259 <LEE>  
A/Cross-references: GB:M31145; NID:g183117; PIDN:AAA52542.1; PID:g183118  
A/Note: part of this sequence, including the amino end of the mature protein, was confirm  
R/Ehrenborg, E.; Larsson, C.; Stern, I.; Janson, M.; Powell, D.R.; Luthman, H.  
Genomics 12, 497-502, 1992  
A/Title: Contiguous localization of the genes encoding human insulin-like growth factor I  
A/Reference number: A42176; PMID:92317971; PMID:1373120  
A/Accession: A42176  
A/Molecule type: DNA  
A/Residues: 174-252, 'W', 254-259 <EHR>  
A/Note: sequence extracted from NCBI backbone (NCBI:95701, NCBI:P.95702)  
R/Luthman, H.; Seoderling-Barrois, J.; Persson, B.; Engberg, C.; Stern, I.; Lake, M.; Frar  
Povoa, G.; Hall, K.; Joernvall, H.  
Eur. J. Biochem. 180, 259-265, 1989  
A/Title: Human insulin-like growth-factor-binding protein. Low-molecular-mass form: prote  
A/Reference number: S03559; PMID:89170723; PMID:2466655  
A/Accession: S43207  
A/Molecule type: mRNA  
A/Residues: 141-259 <LUT1>  
A/Cross-references: EMBL:X15002; NID:g33013; PIDN:CAA33110.1; PID:g33014  
A/Accession: S03559  
A/Molecule type: protein  
A/Residues: 26-259 <LUT2>  
A/Note: 228-Met was also found  
R/Julkinen, M.; Koistinen, R.; Aalto-Setälä, K.; Seppälä, M.; Jaenmäe, O.A.; Kontula,  
FEBS Lett. 236, 295-302, 1988  
A/Title: Primary structure of human insulin-like growth factor-binding protein/placental  
A/Reference number: S01333; PMID:88312985; PMID:2457513  
A/Accession: S01333  
A/Molecule type: mRNA  
A/Residues: 1-252, 'W', 254-259 <JUL>  
A/Cross-references: EMBL:X13405; NID:g35574; PIDN:CAA31771.1; PID:g35575  
R/Brewer, M.T.; Stetler, G.L.; Squires, C.H.; Thompson, R.C.; Busby, W.H.; Clemmons, D.R.  
Biochem. Biophys. Res. Commun. 152, 1289-1297, 1988  
A/Title: Cloning, characterization, and expression of a human insulin-like growth factor

A:Reference number: A27694; MUID:88240345; PMID:2454104  
 A:Accession: A27694  
 A:Molecule type: mRNA  
 A:Residues: 1-54; 'PAAVAVRCAPCLWAPR', 72-212, 'Q', 214-259 <BRE>  
 A:Cross-references: GB:M20841; NID:G183113; PIDN:AA52540.1; PID:G183114  
 A>Note: the authors translated the codon CAG for residue 213 as His; the nucleotide sequ  
 Nucleic Acids Res. 16, 8711, 1988  
 R:Grundmann, U.; Nerlich, C.; Bohm, H.; Rein, T.  
 A:Title: Cloning of cDNA encoding human placental protein 12 (PP12): binding protein for  
 A:Reference number: S01415; MUID:88335621; PMID:3419931  
 A:Accession: S01415  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <GNU>  
 A:Cross-references: EMBL:X12385; NID:G35571; PIDN:CA30942.1; PID:G35572  
 R:Busby Jr., W.H.; Klappper, D.G.; Clemmons, D.R.  
 J. Biol. Chem. 263, 14203-14210, 1988  
 A:Title: Purification of a 31,000-Dalton insulin-like growth factor binding protein from  
 A:Reference number: A30804; MUID:89008261; PMID:2971653  
 A:Accession: A30804  
 A:Molecule type: protein  
 A:Residues: 26-53 <BUS>  
 R:Bell, S.C.; Keyte, J.W.  
 Endocrinology 123, 1202-1204, 1988  
 A:Title: N-terminal amino acid sequence of human pregnancy-associated endometrial alpha-  
 weight IGF binding proteins.  
 A:Reference number: A61099; MUID:88283527; PMID:2456201  
 A:Accession: A61099  
 A:Molecule type: protein  
 A:Residues: 26-40, 'P', 42 <BEL>  
 C:Comment: This protein binds insulin-like growth factor I (somatomedin C) and insulin-I  
 C:Genetics: This protein is not glycosylated.  
 A:Gene: GDB:IGFBP1; IBP1  
 A:Cross-references: GDB:120075; OMIM:146730  
 A:Map position: 7p13-7p12  
 A:Intons: 117/1; 173/3; 216/3  
 C:Superfamily: Insulin-like growth factor binding protein; thyroglobulin type I repeat H  
 C:Keywords: phosphoprotein  
 F:1-25/Domain: signal sequence #status predicted <SIG>  
 F:16-259/Product: insulin-like growth factor-binding protein 1 #status experimental <MUT  
 F:176-251/Domain: thyroglobulin type I repeat homology <TRY1>  
 F:246-248/Region: cell attachment (R-G-D) motif

Query Match 1.9%; Score 5; DB 1; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 LSCRA 168  
 Db 82 LSCRA 86

RESULT 6  
 RNA-binding protein - wood tobacco  
 C:Species: Nicotiana sylvestris (wood tobacco)  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 09-Jul-2004  
 C:Accession: S46286  
 R:Hirose, T.; Sugita, M.; Sugitara, M.  
 Mol. Gen. Genet. 244, 360-366, 1994  
 A:Title: Characterization of a cDNA encoding a novel type of RNA-binding protein in toba  
 A:Reference number: S46286; MUID:94359458; PMID:8078461  
 A:Accession: S46286  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-259 <HIR>  
 A:Cross-references: UNIPROT:Q40436; GB:D26182; NID:G575607; PIDN:BA05170.1; PID:G624929  
 C:Superfamily: wood tobacco RNA-binding protein; ribonucleoprotein repeat homology  
 F:41-108/Domain: ribonucleoprotein repeat homology <RM1>

Query Match 1.9%; Score 5; DB 1; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 132 GSGGG 136  
 Db 129 GSGGG 133

RESULT 7  
 cis-1,2-dihydro-1,2-dihydroxy-naphthalene dehydrogenase (EC 1.3.1.29) - Pseudomonas sp.  
 C:Species: Pseudomonas sp.  
 C:Date: 07-Apr-1994 #sequence\_revision 07-Apr-1994 #text\_change 26-Aug-1999  
 C:Accession: D49343; S27635  
 R:Denome, S.A.; Stanley, D.C.; Olson, E.S.; Young, K.D.  
 J. Bacteriol. 175, 6890-6901, 1993  
 A:Title: Metabolism of dibenzochlorophene and naphthalene in Pseudomonas strains: complete  
 A:Reference number: A49343; MUID:94042852; PMID:8226631  
 A:Accession: D49343  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-259 <DEN>  
 A:Cross-references: GB:M60405  
 C:Genetics:

A:Gene: dox  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
 C:Keywords: oxidoreductase  
 F:5-183/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 1.9%; Score 5; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 129 GSGGS 133  
 Db 147 GSGGS 151

RESULT 8  
 probable short-chain dehydrogenase PA1330 [imported] - Pseudomonas aeruginosa (strain PA  
 F83479  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 09-Jul-2004  
 C:Accession: F83479  
 R:Stover, C.K.; Pham, X.Q.; Ewin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lartig, K.; Llm,  
 .; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337; PMID:10984043  
 A:Accession: F83479  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-259 <STO>  
 A:Cross-references: UNIPROT:Q91414; GB:AE004562; GB:AE004091; NID:G99477263; PIDN:AA60471;  
 C:Genetics: A:Experimental source: strain PA01  
 A:Gene: PA1330  
 C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 1.9%; Score 5; DB 2; Length 259;  
 Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 103 EADEG 107  
 Db 229 EADEG 233

RESULT 9  
 E55117  
 cis-1,2-dihydro-1,2-dihydroxynaphthalene dehydrogenase (EC 1.3.1.29) - Pseudomonas putida  
 N:Alternate names: polycyclic aromatic hydrocarbon dihydrodiol dehydrogenase  
 C:Species: Pseudomonas putida  
 C:Date: 05-May-1995 #sequence\_revision 05-May-1995 #text\_change 09-Jul-2004

C/Accession: E55217  
R/Takizawa, N.; Kaida, N.; Torigoe, S.; Moritani, T.; Sawada, T.; Satoh, S.; Kiyohara, H.  
J. Bacteriol. 176, 2444-2449, 1994  
A/Title: Identification and characterization of genes encoding polycyclic aromatic hydrocarbons  
A/Reference number: A55217; MUID:94209249; PMID:8157615  
A/Accession: E55217  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <TRK>  
A/Cross-references: UNIPROT:Q08669; GB:AB004059; GB:D16629; NID:g2189972; PIDN:BAA20393.  
C/Genetics:  
A/Genes: pahB  
C/Superfamily: ribicol dehydrogenase; short-chain alcohol dehydrogenase homology  
C/Keywords: oxidoreductase  
F/5-183/Domain: short-chain alcohol dehydrogenase homology <SADR>  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 129 GGGGS 133  
Db 147 GGGGS 151  
  
RESULT 10  
B83838  
oxidoreductase BH1506 [imported] - Bacillus halodurans (strain C-125)  
C/Species: Bacillus halodurans  
C/Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 09-Jul-2004  
R/Accession: B83838  
R/Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Saseaki, R.; Masui, N.; Fujii, F.; Hira  
Nucleic Acids Res. 28, 4317-4331, 2000  
A/Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A/Reference number: A83650; MUID:20512582; PMID:11058132  
A/Accession: B83838  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <STO>  
A/Cross-references: UNIPROT:Q9KCR3; GB:AP001512; GB:BA000004; NID:g10174030; PIDN:BAH052  
C/Genetics:  
A/Genes: BH1506  
C/Superfamily: ribicol dehydrogenase; short-chain alcohol dehydrogenase homology  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 186 APRLL 190  
Db 244 APRLL 248  
  
RESULT 11  
C75410  
mechanine aminopeptidase - Deinococcus radiodurans (strain R1)  
C/Species: Deinococcus radiodurans  
C/Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C/Accession: C75410  
R/White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathavan, J.O.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Ma  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A/Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A/Reference number: A75250; MUID:20036896; PMID:10567266  
A/Accession: C75410  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-259 <WHI>  
A/Cross-references: UNIPROT:Q9RUS1; GB:AE001978; GB:AE000513; NID:g6459059; PIDN:AAF1088  
C/Genetics:  
A/Experimental source: strain R1

A/Genes: DR1311  
A/Map position: 1  
C/Superfamily: Escherichia coli methionyl aminopeptidase  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 70 ITADK 74  
Db 233 ITADK 237  
  
RESULT 12  
T28147  
class II histocompatibility antigen M beta chain 2 - chicken  
C/Species: Gallus gallus (chicken)  
C/Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 20-Jun-2000  
C/Accession: T28147  
R/Milne, S.; Kaufman, J.; Beck, S.  
submitted to the EMBL Data Library, May 1998  
A/Description: DNA sequencing and analysis of the chicken major histocompatibility comple  
A/Reference number: Z20475  
A/Accession: T28147  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-259 <MTL>  
A/Cross-references: EMBL:AL023516; PIDN:CAA18967.1  
C/Genetics:  
A/Genes: BMD2  
A/Map position: 116  
A/Intons: 22/1; 116/1; 208/1; 246/1; 256/3  
C/Superfamily: class II histocompatibility antigen; immunoglobulin homology  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 24 ASGDT 28  
Db 160 ASGDT 164  
  
RESULT 13  
I48119  
P-glycoprotein - Chinese hamster (fragment)  
C/Species: Citreolus griseus (Chinese hamster)  
C/Date: 04-Sep-1997 #sequence\_revision 04-Sep-1997 #text\_change 02-Feb-2001  
R/Ng, W.F.; Sarangi, F.; Zastawny, R.L.; Veinot-Drebot, L.; Ling, V.  
Mol. Cell. Biol. 9, 1224-1232, 1989  
A/Title: Identification of members of the P-glycoprotein multigene family.  
A/Reference number: I48119; MUID:89261726; PMID:2566908  
A/Accession: I48119  
A/Status: preliminary; translated from GB/EMBL/DBJ  
A/Molecule type: DNA  
A/Residues: 1-259 <RES>  
A/Cross-references: GB:M25792; NID:g576810; PIDN:AA53439.1; PID:g576812  
C/Genetics:  
A/Genes: pgp3  
A/Intons: 66/3; 142/3; 191/3  
C/Superfamily: multidrug resistance protein; ATP-binding cassette homology  
C/Keywords: ATP; nucleotide binding; P-loop  
F/25-228/Domain: ATP-binding cassette homology <ABC2>  
F/42-49/Region: nucleotide-binding motif A (P-loop)  
  
Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 2 VQLE 6  
Db 2 VQLE 6

Db 52 VQLE 56

## RESULT 14

AC0358  
probable pilli chaperone protein YP02944 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Jul-2004  
C:Accession: AC0358  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tittball, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
Hill, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrall, N.  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AC0358  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-259 <KUR>  
A:Cross-references: UNIPROT:Q82CN7; GB:AL590842; PIDN:CAC92190.1; PID:G15980902; GSPDB:G  
C:Genetics:  
A:Gene: YP02944  
C:Superfamily: chaperone protein papd

Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 216 LTISS 220  
Db 18 LTISS 22

## RESULT 15

D75372  
probable nitrogen regulator - Deinococcus radiodurans (strain R1)  
C:Species: Deinococcus radiodurans  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 09-Jul-2004  
C:Accession: D75372  
R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
M.; Shen, M.; Vamathevan, J.; Lam, P.; McDonald, L.; Uterback, T.; Zalewski, C.; Mc  
S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
Science 286, 1571-1577, 1999  
A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
A:Reference number: A75250; MUID:20036896; PMID:10567266  
A:Accession: D75372  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-259 <WHI>  
A:Cross-references: UNIPROT:Q9RTV7; GB:AE002007; GB:AE000513; NID:g6459402; PIDN:AAF115  
C:Genetics:  
A:Experimental source: strain R1  
A:Gene: DR1646  
A:Map position: 1  
C:Superfamily: regulatory protein fnr; cAMP receptor protein cyclic nucleotide-binding d

Query Match 1.9%; Score 5; DB 2; Length 259;  
Best Local Similarity 100.0%; Pred. No. 7.4e+03;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 GQGTU 123  
Db 132 GQGTU 136

Search completed: September 26, 2005, 13:15:37  
Job time : 41 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 26, 2005, 13:04:41 ; Search time 169 seconds  
(without alignments)  
784.784 Million cell updates/sec

Title: US-09-936-702-3\_COPY\_244\_502

Perfect score: 259  
Sequence: 1 QVQLSGAEVKKPKGSSVKV.....RYTFGGTRLEIKLVRGSG 259

Scoring table: OLIGO  
Gapop 60.0 , Gapext 60.0

Searched: 1613378 seqs, 512079187 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2784

Minimum DB seq length: 259

Maximum DB seq length: 259

Post-processing: Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	6	2.3	259	2 Q8ZT71	Q8ZT71 pyrobaculum
2	6	2.3	259	2 Q9HSW3	Q9HSW3 halobacteri
3	6	2.3	259	2 Q6QHA6	Q6QHA6 homo sapien
4	6	2.3	259	2 Q02049	Q02049 caenorhabdi
5	6	2.3	259	2 Q9V6A4	Q9V6A4 drosophila
6	6	2.3	259	2 Q6QHA4	Q6QHA4 oryza sativ
7	6	2.3	259	2 Q6ZGS2	Q6ZGS2 oryza sativ
8	6	2.3	259	2 Q7XZHO	Q7XZHO oryza sativ
9	6	2.3	259	2 Q851D7	Q851D7 oryza sativ
10	6	2.3	259	2 Q9C8P8	Q9C8P8 arabidopsis
11	6	2.3	259	2 Q9F862	Q9F862 triticum ae
12	6	2.3	259	2 Q9F875	Q9F875 triticum ae
13	6	2.3	259	2 Q66FN3	Q66FN3 yerisina ps
14	6	2.3	259	2 Q6S9V1	Q6S9V1 lepospora
15	6	2.3	259	2 Q7ARN6	Q7ARN6 yerisina ps
16	6	2.3	259	2 Q7B1Z4	Q7B1Z4 yerisina ps
17	6	2.3	259	2 Q7B1Y2	Q7B1Y2 yerisina ps
18	6	2.3	259	2 Q83W92	Q83W92 escherichia
19	6	2.3	259	2 Q7ARKO	Q7ARKO yerisina pe
20	6	2.3	259	2 P74994	P74994 escherichia
21	6	2.3	259	2 Q6MOL1	Q6MOL1 bdellovibri
22	6	2.3	259	2 Q741Z8	Q741Z8 mycobacteri
23	6	1.9	259	1 ALDC_KLSTE	ALDC_KLSTE klebsiella
24	6	1.9	259	1 EXB4_ARATV	EXB4_ARATV arbidopsis
25	6	1.9	259	1 IBP1_HUMAN	IBP1_HUMAN homo sapien
26	6	1.9	259	1 NAH8_PSEPU	NAH8_PSEPU pseudomonas
27	6	1.9	259	1 PYRK_BACAN	PYRK_BACAN bacillus an
28	6	1.9	259	1 PYRK_BACCR	PYRK_BACCR bacillus ce
29	6	1.9	259	1 S5A1_HUMAN	S5A1_HUMAN homo sapien
30	6	1.9	259	1 S5A1_RAT	S5A1_RAT rattus norv
31	6	1.9	259	1 SURE_BACTN	SURE_BACTN bacteroides

32	5	1.9	259	1 UBIE_AGRTS	Q8UHS agrobacteri
33	5	1.9	259	1 YB14_NEIMA	Q9JUV9 neisseria m
34	5	1.9	259	1 YZ38_METUA	Q60293 methanococc
35	5	1.9	259	2 Q648P5	Q648P5 uncultured
36	5	1.9	259	2 Q8PY11	Q8PY11 methanosarc
37	5	1.9	259	2 Q8ZUQ2	Q8ZUQ2 pyrobaculum
38	5	1.9	259	2 Q9HNH8	Q9HNH8 halobacteri
39	5	1.9	259	2 Q96V53	Q96V53 pleurotus o
40	5	1.9	259	2 Q6CKL3	Q6CKL3 kluyveromyc
41	5	1.9	259	2 Q6CKG1	Q6CKG1 kluyveromyc
42	5	1.9	259	2 Q6T413	Q6T413 leposphaer
43	5	1.9	259	2 Q7S5M0	Q7S5M0 neosporea
44	5	1.9	259	2 Q7S669	Q7S669 neosporea
45	5	1.9	259	2 Q96011	Q96011 homo sapien

## ALIGNMENTS

RESULT 1					
ID	Q8ZT71	PRELIMINARY;	PRT;	259 AA.	
AC	Q8ZT71;				
DT	01-MAR-2002 (TREMBlrel. 20, Created)				
DT	01-MAR-2002 (TREMBlrel. 20, Last sequence update)				
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)				
DE	Hypothetical protein PAE3407.				
GN	OrderedlocusNames=PAE3407;				
OS	Pyrobaculum aerophilum.				
OC	Archaea; Crenarchaeota; Thermoprotei; Thermoproteales;				
OC	Thermoproteaceae; Pyrobaculum.				
OX	NCBI_TaxID=13773;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN=IM2 / ATCC 51768 / DSM 7523.				
RX	MEDLINE=21664397; PubMed=11792869; DOI=10.1073/pnas.241636498;				
RA	Fitz-Gibbon S.T., Ladner H., Kim U.-J., Stetter K.O., Simon M.I.,				
RA	Miller J.H.;				
RT	"Genome sequence of the hyperthermophilic crenarchaeon Pyrobaculum				
RT	aerophilum.";				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:984-989(2002).				
DR	EMBL; AE009928; AAL64892.1; -.				
KW	Complete proteome.				
SQ	SEQUENCE 259 AA; 27170 MW; 726890475F69CC00 CRC64;				
Query Match					
Best Local Similarity 100.0%; Pred. No. 4e+03; Length 259;					
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;					
Qy	54 ILDVAN 59				
Db	231 ILDVAN 236				
RESULT 2					
ID	Q9HSW3	PRELIMINARY;	PRT;	259 AA.	
AC	Q9HSW3;				
DT	01-MAR-2001 (TREMBlrel. 16, Created)				
DT	01-MAR-2001 (TREMBlrel. 16, Last sequence update)				
DT	01-JUN-2003 (TREMBlrel. 24, Last annotation update)				
DE	Vng0049h.				
GN	OrderedlocusNames=VNG0049H;				
OS	Halobacterium sp. (strain NRC-1 / ATCC 700922 / JCM 11081).				
OC	Archaea; Euryarchaeota; Halobacteriia; Halobacteriales;				
OC	Halobacteriaceae; Halobacterium.				
OX	NCBI_TaxID=64091;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RX	MEDLINE=20504483; PubMed=11016950; DOI=10.1073/pnas.190337797;				
RA	Ng W.-V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,				
RA	Shukla H.D., Lasky S.R., Baling N.S., Thorsson V., Sbrogna J.,				
RA	Swartzell S., Weir D., Hall J., Dahl T.A., Welti R., Goo Y.A.,				

```

RA Leithauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,
RA Maddocks D.G., Jabloncki P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Beck R.F., Pohlschroder M., Spudich J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ebhardt H., Lowe T.M., Liang P., Riley M., Hood L., DasSarma S.,
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
DR EMBL; AS004973; A018687.1; -.
DR PIR; C84165; C84165.
DR GO; GO:0008757; F:S-adenosylmethionine-dependent methyltransf. . .; IEA.
DR InterPro; IPR006134; Fkm_mtfase.
DR InterPro; IPR000051; SAM_Bind.
DR TrEMBL; TIGR01444; Fkm_fam. 1.
DR Complete proteome.
SQ SEQUENCE 259 AA; 28842 MW; 6D147C0A0CC0784C CRC64;

```

```

Query Match 2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 GSSGVV 20
DB 134 GSSGVV 139

RESULT 3
O6OHA6 PRELIMINARY; PRT; 259 AA.
AC O6OHA6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DE Myoneurin (Fragment).
GN Name=MYNN;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN (1)
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=2034280; PubMed=10873615; DOI=10.1006/dbrc.2000.2862;
RA Alliel P.M., Seddidi N., Goudou D., Cluettens-Diaz C., Romero N.,
RA Velasco E., Rieger F., Perin J.P.,
RT "Myoneurin, a novel member of the BTB/POZ-zinc finger family highly
expressed in human muscle."
RL Biochem. Biophys. Res. Commun. 273:385-391(2000).
RN (2)
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX PubMed=14694499;
RA Cluettens-Diaz C., Bitoun M., Goudou D., Seddidi N., Romero N.,
RA Rieger F., Perin J.P., Alliel P.M.;
RT "Neuromuscular expression of the BTB/POZ and zinc finger protein
myoneurin."
RL Muscle Nerve 29:59-65(2004).
DR EMBL; AY541761; AA587377.1; -.
DR GO; GO:0005515; F:protein binding; IEA.
DR InterPro; IPR000210; BTB_POZ.
DR Pfam; PR00651; BTB_1.
DR SMART; SM00225; BTB_1.
DR PROSITE; PSS0097; BTB_1.
FT NON_TER 1
FT NON_TER 1
SQ SEQUENCE 259 AA; 28992 MW; 556F938E5133EB9F CRC64;

Query Match 2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 143 SELELT 148
DB 202 SELELT 207

```

```

RESULT 4
O02049 PRELIMINARY; PRT; 259 AA.
ID O02049;
AC O02049;
DT 01-JUL-1997 (TrEMBLrel. 04, Created)
DT 01-JUL-1997 (TrEMBLrel. 04, Last sequence update)
DE 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein T20B6.3.
GN Name=T20B6.3; ORFNames=T20B6.3;
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;
OC Rhabditidae; Pelodierinae; Caenorhabditis.
OX NCBI_TaxID=6239;
RN (1)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RX MEDLINE=9089613; PubMed=9851916;
RG Wormbase Consortium;
RT "Genome sequence of the nematode C. elegans: a platform for
investigating biology. The C. elegans Sequencing Consortium."
RL Science 282:2012-2016(1998).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Beck C., Wamsley P.;
RT "The sequence of C. elegans cosmid T20B6."
RL Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN (3)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RA Waterston R.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN (4)
RP SEQUENCE FROM N.A.
RC STRAIN=Bristol N2;
RG Wormbase Consortium;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AF000193; AA52890.1; -.
DR PIR; T15126; T15126.
DR Wormbase; WBGene00020599; T20B6.3.
DR WormPep; T20B6.3; CE13778.
DR InterPro; IPR002952; Eggshell.
DR PRINTS; PR01228; EGGSHL.
KM Hypothetical protein.
SQ SEQUENCE 259 AA; 23054 MW; C5F1690D2CE04CAF CRC64;

```

```

Query Match 2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 131 GSGGCG 136
DB 80 GSGGCG 85

RESULT 5
O9V6A4 PRELIMINARY; PRT; 259 AA.
ID O9V6A4;
AC O9V6A4;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE CG30042-PA.
GN ORFNames=CG30042;
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN (1)
RP SEQUENCE FROM N.A.

```

RX MEDLINE=20196006; PubMed=10711132; DOI=10.1126/science.287.5461.2185;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amaratunga C., Scherer S.E., Li P.W., Hoskins K.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.E., White M.J., Schaefer M., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H., Blazer V.G., Champ M., Pfeiffer B.D.,  
 RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,  
 RA Abil J.F., Agbayanti A., An H.U., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Baou P.V., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
 RA Botkin D., Botchan D.A., Bouck J., Brockstein P., Brodtier P.,  
 RA Butte K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Dey A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Dublin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fowler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck M.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Idegami C.,  
 RA Jaitai M., Kalush F., Kapran G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Liao X.P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Plutman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,  
 RA Shie B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spreading A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodagert, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,  
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
 RL "The genome sequence of *Drosophila melanogaster*.";  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426065; PubMed=12537568;  
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,  
 RA Patel S.E., Adams M., Champ M., Dugan S.P., Frise E., Hodgson A.,  
 RA George R.A., Hoskins R.A., Laverly T., Muzny D.M., Nelson C.R.,  
 RA Pacle J.M., Pask S., Pfeiffer B.D., Richards S., Sodergren E.J.,  
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,  
 RA Weinstein G., Wheeler S.E., Myers E.W., Gibbs R.A., Rubin G.M.,  
 RT "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*  
 RT *melanogaster* euchromatic genome sequence.";  
 RL Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426070; PubMed=12537573;  
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,  
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,  
 RA Ashburner M., Celniker S.E.;  
 RT "The transposable elements of the *Drosophila melanogaster* euchromatin  
 RT a genomics perspective.";  
 RL Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=22426069; PubMed=12537572;  
 RA Maira S., Crosby Y.A., Mungall C.J., Mathews B.B., Campbell K.S.,  
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,  
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Bernan B.P.,  
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Dykdale R.A.,  
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,  
 RA Lewis S.E.;  
 RT "Annotation of the *Drosophila melanogaster* euchromatic genome: a  
 RT systematic review.";  
 RL Genome Biol 3:RESEARCH0083-RESEARCH0083(2002).

RN [5] SEQUENCE FROM N.A.  
 RP FlyBase;  
 RG Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.  
 RL [6]  
 RN RP  
 RP SEQUENCE FROM N.A.  
 RG FlyBase;  
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AE003882; AAFF8554.2; -  
 DR FlyBase; FBgn0050042; CG30042.  
 DR FlyBase; FBgn0050043; CG30043.  
 DR GO: GO:0042302; F:structural constituent of cuticle; IEA  
 DR InterPro: IPR005479; Cph synt h D2.  
 DR InterPro: IPR006189; Insect cuticle.  
 DR Pfam: PF00379; Chitin bind\_4; 1.  
 DR PRINTS: PR00947; CUTICLE.  
 DR PROSITE: PS00867; GPSAE\_2; UNKNOWN\_1.  
 DR PROSITE; PS00233; CUTICLE; UNKNOWN\_1.  
 QO SEQUENCE 259 AA; 26957 MW; ED2A2B6758EF1FD0 CRC64;

```

Query Match      2.3%; Score 6; DB 2; Length 259;
Best Local Similarity 100.0%; Pred. No. 4e+03;
Matches      6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy      129 GGGGSG 134
          |||||
db      105 GGGGSG 110

```

RESULT 6	
Q6QHA4	
ID Q6QHA4	PRELIMINARY;
SECURITY	PRT; 259 AA

DT 05-JUL-2004 (TREMBLrel. 27, Created)  
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)  
DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
DE Glutathione S-transferase.  
GN Name=GSTJ18; Synonyms=P0425P05.18;  
OS Oryza sativa (japonica cultivar-group).  
OC Eukaryota; Viridiplantae; Streptophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriactaceae; Oryzaceae; Oryza.  
NCBI\_TaxID=39947;

RP SEQUENCE FROM N.A.  
RX PubMed=15069639; DOI=10.1007/s00438-004-1006-8;  
RA Soranzo N., Sari Gorla M., Mizzi L., De Toma G., Froya C.,  
RT "organisation and structural evolution of the rice glutathione S-  
transferase gene family."; *PL*  
PL Mol. Genet. Genomics 271:511-521(2004).

RP SEQUENCE FROM N.A.  
RA Saeki T., Matsumoto K.;  
RT "Oryza sativa hypnobarate (GA3) genomic DNA, chromosome 6, PAC  
clone:IP0425F05.-";  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY541763; AAC93256.1; -  
DR EMBL: AP003569; BA037467.1; -  
DR GO: 0016740; F:transferrase activity; IEA.  
DR InterPro: IPR010987; GST\_C like.  
DR InterPro: IPR004045; GST\_Nterm.  
DR Pfam: PF02798; GST\_N, 1.  
KM Transferrase.  
SQ SEQUENCE 259 AA; 27478 MW; 526A5DEFD7C0F3DB4 CRC64;

Query Match	2.3%	Score 6	DB 2	Length 259
Best Local Similarity	100.0%	Pred. No. 4e+03		
Matches	6	Conservative	0	Mismatches 0
				Indels 0
				Gaps 0
Qy	185	QAPRLL	190	
Db	101	QAPRLL	106	

```

RESULT 7
Q6ZGS2 PRELIMINARY; PRT; 259 AA.
AC Q6ZGS2;
DT 05-JUL-2004 (TEMBLrel. 27, Created)
DT 05-JUL-2004 (TEMBLrel. 27, Last sequence update)
DE Plant disease resistance polypeptide-like.
Name=OJ1148.D05.10;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Sasakita T., Matsumoto T., Yamamoto K.;
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP004118; BAD07691.1; -.
KW Polypeptide.
SQ SEQUENCE 259 AA; 27424 MW; F380F493462F69DA CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 132 GSGGGG 137
Db 60 GSGGGG 65

RESULT 8
Q7XZH0 PRELIMINARY; PRT; 259 AA.
ID Q7XZH0;
DT 01-OCT-2003 (TEMBLrel. 25, Created)
DT 01-OCT-2003 (TEMBLrel. 25, Last sequence update)
DT 25-OCT-2004 (TEMBLrel. 28, Last annotation update)
DE Hypothetical protein OSJNB0033J23.5 (Hypothetical protein
OSJNB0028F23.3).
GN Name=OSJNB0033J23.5; Synonyms=OSJNB0028F23.3;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sub B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL "Oryza sativa chromosome 3 BAC OSJNB0028F23 genomic sequence.";
RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Buell R.;

```

```

RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC137507; AAP73847.1; -.
DR EMBL; AC135595; AAT77902.1; -.
DR Gramene; Q7XZH0; -.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 28114 MW; 1805F80A7CEBD524 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 130 GSGGGG 135
Db 158 GSGGGG 163

RESULT 9
Q851D7 PRELIMINARY; PRT; 259 AA.
ID Q851D7;
AC Q851D7;
DT 01-JUN-2003 (TEMBLrel. 24, Created)
DT 01-JUN-2003 (TEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TEMBLrel. 26, Last annotation update)
DE Hypothetical protein OSJNB0021P10.5.
GN Name=OSJNB0021P10.5;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzaceae; Oryza.
NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltzin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padrosh D.W., Tallon L.J., Koo H., Ziemann V., Heiao J., Blunt S.,
RA Vanaken S.S., Riedmuller S.B., Uteback T.T., Feldlyum T.V.,
RA Yang Q.Q., Haas B.J., Sun B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC123974; AA019373.1; -.
DR Gramene; Q851D7; -.
DR InterPro; IPR010847; Hml.
DR Pfam; PF07320; Hml; 1.
KW Hypothetical protein.
SQ SEQUENCE 259 AA; 27972 MW; 03F79FD4DCE2E701 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 126 VTSGGG 131
Db 138 VTSGGG 143

RESULT 10
Q9C8P8 PRELIMINARY; PRT; 259 AA.
ID Q9C8P8;
AC Q9C8P8;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TEMBLrel. 27, Last annotation update)
DE Helix-loop-helix protein 1A, putative; 28707-26892 (At1g35460/F12A4_2)
DE [putative DHL transcription factor].
GN Name=F12A4_2;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsie.

```

OX NCBI\_TaxID=3702;  
 RN  
 RP SEQUENCE FROM N.A.  
 RA Iln X., Kaul S., Town C.D., Benito M., Creasy T.H., Haas B.J., Wu D.,  
 RA Maiti R., Roming C.M., Koo H., Fujii C.Y., Uterback T.R.,  
 RA Barnstead M.E., Bowman C.J., White O., Nieman W.C., Fraser C.M.;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Town C.D., Kaul S.;  
 RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shin P., Banh J.,  
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,  
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,  
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,  
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,  
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,  
 RA Ecker J.R.;  
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RA Shin P., Chen H., Cheuk R., Kim C.J., Koesema E., Meyers M.C.,  
 RA Banh J., Bowser L., Carninci P., Dale J.M., Goldsmith A.D.,  
 RA Hayashizaki Y., Ishida J., Jiang P.X., Jones T., Kamiya A.,  
 RA Karlin-Neumann G., Kawai J., Lam B., Lee J.M., Lin J., Liu S.X.,  
 RA Miranda M., Narusaka M., Nguyen M., Onodera C.S., Palm C.J.,  
 RA Pham P.K., Quach H.L., Sakurai T., Satou M., Seki M., Southwick A.,  
 RA Tang C.C., Toriumi M., Yamada K., Yamamura Y., Yu G., Yu S.,  
 RA Shinozaki K., Davis R.W., Theologis A., Ecker J.R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [5]  
 RP SEQUENCE FROM N.A.  
 RA TISSUE=Flower;  
 RC MEDLINE=2598051; PubMed=12679534; DOI=10.1093/molbev/meg088;  
 RX Heim M.A., Jakoby M., Werber M., Martin C., Weisshaar B., Bailey P.C.;  
 RA "The basic helix-loop-helix transcription factor family in plants: a  
 RT genome-wide study of protein structure and functional diversity.";  
 RL Mol. Biol. Evol. 20:735-747(2003).  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Flower;  
 RA Jakoby M.J., Heim M.A., Bailey P., Martin C., Weisshaar B.;  
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: Contains 1 basic helix-loop-helix (bHLH) domain.  
 DR EMBL; AC023064; AAG52112.1; -;  
 DR EMBL; AY093794; AAM10410.1; -;  
 DR EMBL; AF411791; AAL06481.1; -;  
 DR EMBL; AF488612; AAM10958.1; -;  
 DR PIR; F86475; F86475. -;  
 DR InterPro; IPR001092; HLH\_basic.  
 DR Pfam; PF00010; HLH; 1.  
 DR SMART; SMO0353; HLH; 1.  
 DR PROSITE; PSS0888; HLH; 1.  
 SQ SEQUENCE 259 AA; 28526 MW; D4B99C758208C1D4 CRC64;  
 QY 206 SSGSGG 211  
 DB 107 SSGSGG 112  
 RESULT 11  
 ID Q9FS62 PRELIMINARY; PRT; 259 AA.  
 AC Q9FS62;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)

DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Gamma-gliadin (Fragment).  
 GN Name=GAG56D;  
 OS Triticum aestivum subsp. spelta.  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Triticaceae; Triticum.  
 OX NCBI\_TaxID=58933;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA von Bueren M.;  
 RL "Polymorphisms in two homeologous gamma-gliadin genes and the  
 RT evolution of cultivated wheat.";  
 RL Genet. Resour. Crop Evol. 0:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA von Bueren M.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ389674; CAC11087.1; -;  
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR001376; Gliadin.  
 DR InterPro; IPR001954; Gliadin.  
 DR Pfam; PF00234; TRY\_alpha\_amyl; 1.  
 DR PRINTS; PR00208; GLIADGLUTEN.  
 DR SMART; SMO0499; AAI; 1.  
 FT NON\_TER 1 259  
 FT TER 1 259  
 SQ SEQUENCE 259 AA; 29961 MW; 89A3073443C4191A CRC64;  
 QY 119 GQGTIV 124  
 DB 215 GQGTIV 220  
 RESULT 12  
 ID Q9FS75 PRELIMINARY; PRT; 259 AA.  
 AC Q9FS75;  
 DT 01-MAR-2001 (TREMBlrel. 16, Created)  
 DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
 DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)  
 DE Gamma-gliadin (Fragment).  
 GN Name=GAG56D;  
 OS Triticum aestivum (Wheat).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae;  
 OC Triticaceae; Triticum.  
 OX NCBI\_TaxID=4565;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA von Bueren M.;  
 RL "Polymorphisms in two homeologous gamma-gliadin genes and the  
 RT evolution of cultivated wheat.";  
 RL Genet. Resour. Crop Evol. 0:0-0(0).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA von Bueren M.;  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AJ389668; CAC11057.1; -;  
 DR GO; GO:0045735; F:nutrient reservoir activity; IEA.  
 DR InterPro; IPR003612; AAI.  
 DR InterPro; IPR001376; Gliadin.  
 DR InterPro; IPR001954; Gliadin.  
 DR Pfam; PF00234; TRY\_alpha\_amyl; 1.  
 DR PRINTS; PR00208; GLIADGLUTEN.  
 DR SMART; SMO0499; AAI; 1.  
 DR PRINTS; PR00209; GLIADIN.  
 DR SMART; SMO0499; AAI; 1.

```

FT NON_TER 1 1
FT NON_TER 259 259
SQ SEQUENCE 259 AA; 30020 MW; BECF13EB3B0812B9 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 119 GGGTLV 124
Db 216 GGGTLV 221

RESULT 13
Q66FN3 PRELIMINARY; PRT; 259 AA.
AC O66FN3;
DT 25-OCT-2004 (TReMBLrel. 28, Last sequence update)
DE Putative IS100 transposase.
GN ORFNames=YPTB0303, YPTB1133, YPTB1586, YPTB1809, YPTB3876;
OS Yersinia pseudotuberculosis IP 32953.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=273123;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=IP 32953;
RX PubMed=1535858;
RA Chain P.S.G., Carniel E., Larimer F.W., Lamerdin J., Stoutland P.O.,
RA Regala W.M., Georgescu A.M., Verges L.W., Land M.L., Motin L.V.,
RA Brubaker R.R., Fowler J., Hinebusch B.J., Marceau M., Medigue C.,
RA Simonet M., Cheneal-Francois V., Souza B., Dacheux D., Elliott J.M.,
RA Derbise A., Hauser L.J., Garcia E.;
RT "Insights into the genome evolution of Yersinia pestis through whole
RT genome comparison with Yersinia pseudotuberculosis.";
RL Proc. Natl. Acad. Sci. U.S.A. 101:13826-13831(2004).
DR EMBL; BX936398; CAH19543.1; -.
DR EMBL; BX936398; CAH20373.1; -.
DR EMBL; BX936398; CAH20825.1; -.
DR EMBL; BX936398; CAH21048.1; -.
DR EMBL; BX936398; CAH23114.1; -.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001270; Chaprin_c1pA/B.
DR Pfam; PFO1695; IseB_1.
DR PRINTS; PR00300; CLPROTEASEA.
DR SMART; SM00382; AAA, 1.
KW ATP-binding.
SQ SEQUENCE 259 AA; 29295 MW; E72FD9033B80E379 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 APRLLI 191
Db 164 APRLLI 169

RESULT 14
Q6S9V1 PRELIMINARY; PRT; 259 AA.
AC O6S9V1;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE LipL32 (Fragment).
GN Name=LipL32;
OS Leptospira borgpetersenii.
OC Bacteria; Spirochaetes; Spirochaetales; Leptospiraceae; Leptospira.
OX NCBI_TaxID=174;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Mus 127;
RX PubMed=15090524; DOI=10.1128/JB.186.9.2818-2828.2004;
RA Haake D.A., Suard M.A., Kelley M.M., Dundoo M., Alt D.P.,
RA Zuercher R.L.;
RT "Molecular evolution and mosaicism of leptospiral outer membrane
RT proteins involves horizontal DNA transfer.";
RL J. Bacteriol. 186:2818-2828(2004).
DR EMBL; AY461894; AAS21764.1; -.
FT NON_TER 1 1
FT NON_TER 259 259
SQ SEQUENCE 259 AA; 28210 MW; A845A8035FE5A463 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 182 KPGQAP 187
Db 63 KPGQAP 68

RESULT 15
Q7ARN6 PRELIMINARY; PRT; 259 AA.
AC Q7ARN6;
DT 05-JUL-2004 (TReMBLrel. 27, Created)
DT 05-JUL-2004 (TReMBLrel. 27, Last sequence update)
DE ORF1 (ORF 77) (ORF; putative transposase).
OS Yersinia pestis.
OG Plasmid 9.5 kbp pesticin.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
OC Enterobacteriaceae; Yersinia.
OX NCBI_TaxID=632;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=6/69;
RA Buchrieser C., Ruenloek C., Couve E., Frangeul L., Billault A.,
RA Kunst F., Carniel E., Glaser P.;
RL Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=EV76-6; PLASMID=9.5 kbp pesticin;
RX MEDLINE=97221614; PubMed=9068660;
RA McDonough K.A., Hare J.M.;
RT "Homology with a repeated Yersinia pestis DNA sequence IS100
RT correlates with pesticin sensitivity in Yersinia pseudotuberculosis.";
RL J. Bacteriol. 179:2081-2085(1997).
DR EMBL; AL031866; CAA21334.1; -.
DR EMBL; AL031866; CAA21402.1; -.
DR EMBL; U59875; AAC44982.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:000166; F:nucleotide binding; IEA.
DR InterPro; IPR003593; AAA_ATPase.
DR InterPro; IPR001270; Chaprin_c1pA/B.
DR Pfam; PFO1695; IseB_1.
DR PRINTS; PR00300; CLPROTEASEA.
DR SMART; SM00382; AAA, 1.
KW ATP-binding; Plasmid.
SQ SEQUENCE 259 AA; 29295 MW; E72FD9033B80E379 CRC64;

Query Match
Best Local Similarity 100.0%; Score 6; DB 2; Length 259;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 186 APRLLI 191
Db 164 APRLLI 169

```

Wed Sep 28 10:54:06 2005

us-09-936-702-3\_copy\_244\_502.olgsiz.rup

Page 7

Search completed: September 26, 2005, 13:14:50  
Job time : 170 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 10:28:40 ; Search time 0.001 Seconds  
(without alignments)  
548.562 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPKGSSVKY.....RTFGQTRLEIKLVPRGSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapext 0.5

Searched: 27 seqs, 2118 residues

Total number of hits satisfying chosen parameters: 27

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 27 summaries

Database : rag.subdb.\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES				Description
Result No.	Score	Query Length	ID	
1	88	34.0	95 1 ABO27152	Human germline I1g
2	88	34.0	95 1 ABO27151	Human germline I1g
3	88	34.0	95 1 ADC99850	Germline VK gene L
4	88	34.0	95 1 ADC99846	Germline VK gene L
5	88	34.0	95 1 ADC99838	Germline VK gene L
6	88	34.0	95 1 ADD05454	Anti-MUC18 antibod
7	88	34.0	95 1 ADD05450	Anti-MUC18 antibod
8	88	34.0	95 1 ADP10095	VEGF antibody I1gh
9	88	34.0	95 1 ADP10197	Antibody light cha
10	88	34.0	95 1 ADP10198	Antibody light cha
11	88	34.0	95 1 ADP09990	Antibody light cha
12	88	34.0	95 1 ADF09991	Antibody light cha
13	88	34.0	95 1 ADF10096	VEGF antibody I1gh
14	88	34.0	95 1 ADF09880	Anti-MUC18 monoclo
15	88	34.0	95 1 ADF09888	Anti-MUC18 monoclo
16	88	34.0	95 1 ADF09892	Anti-MUC18 monoclo
17	88	34.0	95 1 ADJ80271	Vkappa gene locus
18	88	34.0	95 1 ADJ80272	Vkappa gene locus
19	88	34.0	95 1 ADO07341	Human antibody L16
20	88	34.0	95 1 ADO07340	Human antibody L2
21	30	11.6	32 1 ADR43440	FRL variant #11.
22	30	11.6	32 1 ADR31585	Murine TBS-C21 ant
23	29	11.2	32 1 ADR43442	FRL variant #13.
24	29	11.2	32 1 ADR43441	FRL variant #12.
25	29	11.2	32 1 ADR31587	Murine TBS-C21 ant
26	29	11.2	32 1 ADR31586	Murine TBS-C21 ant
27	24	9.3	26 1 AAU90894	Insulin/insulin-I1

ALIGNMENTS

RESULT 1

ABO27152	ID	ABO27152 standard; protein; 95 AA.
XX	AC	ABO27152;
XX	DT	10-SEP-2003 (first entry)
XX	DE	Human germline light chain variable region gene segment #32.
XX	KW	Human: light chain variable region; VK: humanised antibody;
XX	KW	chimeric antibody; complementarity determining region; CDR;
XX	XX	canonical CDR structure type.
XX	OS	Homo sapiens.
XX	PN	US2003039649-A1.
XX	PD	27-FEB-2003.
XX	PF	12-JUL-2002; 2002US-00194975.
XX	PR	12-JUL-2001; 2001US-030511P.
XX	PA	(FOOT/) FOOT J.
XX	PI	Foot J.
XX	DR	WPI; 2003-492151/46.
XX	PT	Making humanized antibody for converting antibody, by making chimeric
XX	PT	antibodies containing complementarity determining region from non-human
XX	XX	antibody and appropriate framework sequences of human antibodies.
XX	PS	Example 1; Fig 2; 31pp; English.
XX	CC	The invention describes a method of making a humanised antibody,
XX	CC	comprising making chimeric antibodies containing a complementarity
XX	CC	determining region (CDR) from a non-human antibody and appropriate
XX	CC	framework sequences (I) of human antibodies. (I) is selected by using
XX	CC	canonical CDR structure types of non-human antibody in comparison to
XX	CC	germline canonical CDR structure types of human antibodies as the basis
XX	CC	for selection, for humanisation. The method is useful for making a
XX	CC	humanised antibody or a converted antibody. The method is applicable for
XX	CC	converting a subject antibody sequence of any subject species to a less
XX	CC	immunogenic form suitable for use in an object species. The method is
XX	CC	reliable for identifying suitable human framework sequences to support
XX	CC	non-human CDR regions and to provide humanised antibodies that retain
XX	CC	high antigen binding with low immunogenicity in humans, without the need
XX	CC	for direct comparison of framework sequences, without the need for
XX	CC	determining critically important amino acid residues in the framework,
XX	CC	and without the need for multiple iteration and construction to obtain
XX	CC	humanised antibodies with suitable therapeutic properties. The antibody
XX	CC	has high affinity and low immunogenicity without need for comparing
XX	CC	framework sequences between non-human and human antibodies. This sequence
XX	CC	represents a human light chain variable region gene segment used in the
XX	CC	creation of humanised antibodies
XX	XX	Sequence 95 AA:
XX	SO	Query Match 34.0%; Score 88; DB 1; Length 95;
XX	SO	Best Local Similarity 92.6%; Pred. No. 6.6;
XX	SO	Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
Qy	387	EELETSPTLTSVPEERATLSGRASESVSLDAWYQGRGAPRLTIYGASRARGVA 446
Db	1	EIVTQSPPTLTVSPERATLSGRASQSVSNLAWYQGRGAPRLTIYGASRARGVA 60
Qy	447	RFGSGSGAEFTLTISLQSEDPFAVYYCOQYNNWP 481
Db	61	RFGSGSGAEFTLTISLQSEDPFAVYYCOQYNNWP 95

RESULT 2

AB027151  
ID ABO27151 standard; protein; 95 AA.  
XX  
AC ABO27151;  
XX  
DT 10-SEP-2003 (first entry)  
XX  
DE Human germline light chain variable region gene segment #31.  
XX  
KW Human; light chain variable region; VK; humanised antibody;  
KM chimeric antibody; complementarity determining region; CDR;  
KW canonical CDR structure type.  
XX  
OS Homo sapiens.  
PN US2003039649-A1.  
XX  
PD 27-FEB-2003.  
XX  
PF 12-JUL-2002; 2002US-00194975.  
XX  
PR 12-JUL-2001; 2001US-0305111P.  
XX  
PA (FOOT/) FOOT/ J.  
XX  
PI Foote J;  
XX  
DR WPI; 2003-492151/46.  
XX  
PT Making humanized antibody for converting antibody, by making chimeric  
PT antibodies containing complementarity determining region from non-human  
PT antibody and appropriate framework sequences of human antibodies.  
PS  
PS Example 1; Fig 2; 31pp; English.  
XX  
XX The invention describes a method of making a humanised antibody,  
CC comprising making chimeric antibodies containing a complementarity  
CC determining region (CDR) from a non-human antibody and appropriate  
CC framework sequences (I) of human antibodies. (I) is selected by using  
CC canonical CDR structure types of non-human antibody in comparison to  
CC germline canonical CDR structure types of human antibodies as the basis  
CC for selection, for humanisation. The method is useful for making a  
CC humanised antibody or a converted antibody. The method is applicable for  
CC converting a subject antibody sequence of any subject species to a less  
CC immunogenic form suitable for use in an object species. The method is  
CC reliable for identifying suitable human framework sequences to support  
CC non-human CDR regions and to provide humanised antibodies that retain  
CC high antigen binding with low immunogenicity in humans, without the need  
CC for direct comparison of framework sequences, without the need for  
CC determining critically important amino acid residues in the framework,  
CC and without the need for multiple iteration and construction to obtain  
CC humanised antibodies with suitable therapeutic properties. The antibody  
CC has high affinity and low immunogenicity without need for comparing  
CC framework sequences between non-human and human antibodies. This sequence  
CC represents a human light chain variable region gene segment used in the  
CC creation of humanised antibodies  
XX  
SQ Sequence 95 AA;  
XX  
Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 387 ELETQSPATLVSPEERATLSCRASSEVSDDLAWYQKQKGPAPRLIYGASTRATGVA 446  
Db 1 EIVWTQSPATLVSPEERATLSCRASQSVSSNLAWYQKQKGPAPRLIYGASTRATGIPA 60  
QY 447 RFGSGSGAEFTLTITSSLOSEDPFVYVYCCQYNNWP 481  
Db 61 RFGSGSGAEFTLTITSSLOSEDPFVYVYCCQYNNWP 95  
RESULT 3

ADC99850  
ID ADC99850 standard; protein; 95 AA.  
XX  
AC ADC99850;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Germline VK gene I2 region protein SEQ ID 79.  
XX  
XX anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;  
KM cytotoxic; melanoma; oesophageal; pancreatic; colorectal tumour;  
KM cervical carcinoma; intraepithelial neoplasia; colorectal; breast;  
KM lung cancer; germline VK region.  
XX  
OS Unidentified.  
PN WO2003057838-A2.  
XX  
PD 17-JUL-2003.  
XX  
PF 26-DEC-2002; 2002WO-US041581.  
XX  
PR 28-DEC-2001; 2001US-0346299P.  
XX  
PA (ABGE-) ABGENIX INC.  
XX  
PI Gudae J;  
XX  
DR WPI; 2003-587113/55.  
XX  
PT New human anti-MUC18 monoclonal antibodies, useful for treating a disease  
PT or condition associated with expression of MUC18 in a patient, e.g.  
PT tumors, cancers, and other malignancies.  
PS  
PS Example 2; SEQ ID NO 79; 78pp; English.  
XX  
XX The invention relates to a novel isolated monoclonal antibody comprising  
CC a heavy or light chain amino acid or a heavy or light chain variable  
CC domain where the antibody binds to MUC18. The monoclonal antibody of the  
CC invention demonstrates cytotoxic activity and may be useful for treating  
CC a disease or condition associated with the expression of MUC18 on the  
CC cell surface such as tumours, specifically melanoma, oesophageal,  
CC pancreatic or colorectal tumours, carcinomas, particularly cervical  
CC carcinomas and cervical intraepithelial neoplasia and cancers including  
CC colorectal, breast or lung cancer, as well as other malignancies. The  
CC current sequence is that of the germline VK gene region protein of the  
CC invention used to analyse the anti-human MUC18 monoclonal antibody  
CC sequences.  
XX  
SQ Sequence 95 AA;  
XX  
Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
QY 387 ELETQSPATLVSPEERATLSCRASSEVSDDLAWYQKQKGPAPRLIYGASTRATGVA 446  
Db 1 EIVWTQSPATLVSPEERATLSCRASQSVSSNLAWYQKQKGPAPRLIYGASTRATGIPA 60  
QY 447 RFGSGSGAEFTLTITSSLOSEDPFVYVYCCQYNNWP 481  
Db 61 RFGSGSGAEFTLTITSSLOSEDPFVYVYCCQYNNWP 95  
RESULT 4  
ID ADC99846 standard; protein; 95 AA.  
XX  
AC ADC99846;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Germline VK gene I2 region protein SEQ ID 75.

XX anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;  
 KM cytosolic; melanoma; oesophageal; pancreatic; colorectal tumour;  
 KM cervical carcinoma; intraepithelial neoplasia; colorectal; breast;  
 KM lung cancer; germine VK region.  
 XX Unidentified.  
 OS  
 XX WO2003057838-A2.  
 PN  
 XX 17-JUL-2003.  
 PD  
 XX 26-DEC-2002; 2002WO-US041581.  
 PF  
 XX 28-DEC-2001; 2001US-0346299P.  
 PR  
 XX (ABGE-) ABGENIX INC.  
 PA  
 XX Gudae J;  
 PI  
 XX WPI; 2003-587113/55.  
 DR  
 XX  
 PT New human anti-MUC18 monoclonal antibodies, useful for treating a disease  
 PT or condition associated with expression of MUC18 in a patient, e.g.  
 PT tumors, cancers, and other malignancies.  
 PS  
 XX Example 2; SEQ ID NO 75; 78pp; English.  
 CC The invention relates to a novel isolated monoclonal antibody comprising  
 CC a heavy or light chain amino acid or a heavy or light chain variable  
 CC domain where the antibody binds to MUC18. The monoclonal antibody of the  
 CC invention demonstrates cyostatic activity and may be useful for treating  
 CC a disease or condition associated with the expression of MUC18 on the  
 CC cell surface such as tumours, specifically melanoma, oesophageal,  
 CC pancreatic or colorectal tumours, carcinomas, particularly cervical  
 CC carcinomas and cervical intraepithelial neoplasia and cancers including  
 CC colorectal, breast or lung cancer, as well as other malignancies. The  
 CC current sequence is that of the germine VK gene region protein of the  
 CC invention used to analyse the anti-human MUC18 monoclonal antibody  
 CC sequences.  
 CC  
 CC Sequence 95 AA;  
 SQ  
 XX  
 XX Query Match 34.0%; Score 88; DB 1; Length 95;  
 XX Best local Similarity 92.6%; Pred. No. 6.6;  
 XX Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTGSPATLSVSPGERATLSCRASESVSDDLAWYQKRGQAPRLLIYGASTRATGVA 446  
 DB 1 EIVMTGSPATLSVSPGERATLSCRAQSQSVSNLAWYQKRGQAPRLLIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTSSLSQSEDPFAYVYCCQYNNMP 481  
 DB 61 RFSGSGSGTEFTLTSSLSQSEDPFAYVYCCQYNNMP 95

RESULT 5  
 ADC99838  
 ID ADC99838 standard; protein; 95 AA.  
 AC  
 XX ADC99838;  
 AC  
 XX 01-JAN-2004 (first entry)  
 DT  
 XX  
 XX Germine VK gene L2 region protein SEQ ID 67.  
 DE  
 XX anti-human MUC18 monoclonal antibody; heavy; light chain variable domain;  
 KM cytosolic; melanoma; oesophageal; pancreatic; colorectal tumour;  
 KM cervical carcinoma; intraepithelial neoplasia; colorectal; breast;  
 KM lung cancer; germine VK region.  
 XX Unidentified.  
 OS  
 XX

PN WO2003057838-A2.  
 PD 17-JUL-2003.  
 PF 26-DEC-2002; 2002WO-US041581.  
 PR 28-DEC-2001; 2001US-0346299P.  
 PA (ABGE-) ABGENIX INC.  
 PI Gudae J;  
 PI WPI; 2003-587113/55.  
 DR  
 XX  
 PT New human anti-MUC18 monoclonal antibodies, useful for treating a disease  
 PT or condition associated with expression of MUC18 in a patient, e.g.  
 PT tumors, cancers, and other malignancies.  
 PS  
 XX Example 2; SEQ ID NO 67; 78pp; English.  
 CC The invention relates to a novel isolated monoclonal antibody comprising  
 CC a heavy or light chain amino acid or a heavy or light chain variable  
 CC domain where the antibody binds to MUC18. The monoclonal antibody of the  
 CC invention demonstrates cyostatic activity and may be useful for treating  
 CC a disease or condition associated with the expression of MUC18 on the  
 CC cell surface such as tumours, specifically melanoma, oesophageal,  
 CC pancreatic or colorectal tumours, carcinomas, particularly cervical  
 CC carcinomas and cervical intraepithelial neoplasia and cancers including  
 CC colorectal, breast or lung cancer, as well as other malignancies. The  
 CC current sequence is that of the germine VK gene region protein of the  
 CC invention used to analyse the anti-human MUC18 monoclonal antibody  
 CC sequences.  
 CC  
 CC Sequence 95 AA;  
 SQ  
 XX  
 XX Query Match 34.0%; Score 88; DB 1; Length 95;  
 XX Best local Similarity 92.6%; Pred. No. 6.6;  
 XX Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTGSPATLSVSPGERATLSCRASESVSDDLAWYQKRGQAPRLLIYGASTRATGVA 446  
 DB 1 EIVMTGSPATLSVSPGERATLSCRAQSQSVSNLAWYQKRGQAPRLLIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTSSLSQSEDPFAYVYCCQYNNMP 481  
 DB 61 RFSGSGSGTEFTLTSSLSQSEDPFAYVYCCQYNNMP 95

RESULT 6  
 ADD05454  
 ID ADD05454 standard; protein; 95 AA.  
 AC  
 XX ADD05454;  
 AC  
 XX 01-JAN-2004 (first entry)  
 DT  
 XX  
 XX Anti-MUC18 antibody light chain variable region L2 protein, SEQ ID 79.  
 DE  
 XX monoclonal antibody; tumour; MUC18; proliferation; cytostatic; vaccine;  
 KM antigen; tumour metastasis; melanoma; metastatic; human; light chain.  
 XX Unidentified.  
 OS  
 XX WO2003057006-A2.  
 PN  
 XX 17-JUL-2003.  
 PD  
 XX 26-DEC-2002; 2002WO-US041582.  
 PF  
 XX 28-DEC-2001; 2001US-0346460P.  
 PR  
 XX (ABGE-) ABGENIX INC.  
 PA  
 XX

PI Gudaa J, Bar-Eli M;  
XX  
XX WPI; 2003-577496/54.  
XX  
PT Use of monoclonal antibodies against MUC18 antigen, for diagnosing and  
PT treating tumors, inhibiting tumor growth, inhibiting cell invasion  
PT associated with melanoma, or increasing survival of an animal having a  
PT metastatic tumor.  
XX  
XX Disclosure; SEQ ID NO 79; 87pp; English.  
XX  
CC The invention relates to a novel monoclonal antibody used for inhibiting  
CC tumour growth in an animal. The tumour inhibition process comprises  
CC selecting an animal in need of treatment for a tumour, providing a  
CC monoclonal antibody comprising a heavy chain amino acid, where the  
CC antibody consists of any one of 10 fully defined sequences of 117-123  
CC amino acids given in the specification, and where the monoclonal antibody  
CC binds MUC18, and contacting the tumour with the antibody resulting in  
CC inhibited proliferation of the cells. The monoclonal antibody has  
CC cytostatic and can be used in the production of a vaccine. The monoclonal  
CC antibodies against the MUC18 antigen are useful for diagnosing and  
CC treating tumours, inhibiting tumour growth (e.g. melanoma, lung tumour or  
CC tumour metastasis), inhibiting cell invasion associated with melanoma, or  
CC increasing survival of an animal having a metastatic tumour. This  
CC sequence represents an anti-MUC18 antibody light chain, variable region,  
CC protein of the invention.  
XX  
SQ Sequence 95 AA;  
XX  
Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
XX  
QY 387 ELELTOSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVPA 446  
Db 1 EIVMTQSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVPA 60  
XX  
QY 447 RFGSGSGAEFTLTITSSLSQSEDPFAVYCCQYNNWP 481  
Db 61 RFGSGSGTEFTLTITSSLSQSEDPFAVYCCQYNNWP 95  
XX  
RESULT 7  
ADD05450  
ID ADD05450 standard; protein; 95 AA.  
XX  
AC ADD05450;  
XX  
DT 01-JAN-2004 (first entry)  
XX  
DE Anti-MUC18 antibody light chain variable region L2 protein, SEQ ID 75.  
XX  
KM monoclonal antibody; tumour; MUC18; proliferation; cytostatic; vaccine;  
KM antigen; tumour metastasis; melanoma; metastatic; human; light chain.  
XX  
OS Unidentified.  
XX  
PN WO2003057006-A2.  
XX  
PD 17-JUL-2003.  
XX  
PF 26-DEC-2002; 2002WO-US041582.  
XX  
PR 28-DEC-2001; 2001US-0346460P.  
XX  
PA (ABGE-) ABGENIX INC.  
XX  
PI Gudaa J, Bar-Eli M;  
XX  
XX WPI; 2003-577496/54.  
XX  
PT Use of monoclonal antibodies against MUC18 antigen, for diagnosing and  
PT treating tumors, inhibiting tumor growth, inhibiting cell invasion

PT associated with melanoma, or increasing survival of an animal having a  
PT metastatic tumor.  
XX  
XX Disclosure; SEQ ID NO 75; 87pp; English.  
XX  
CC The invention relates to a novel monoclonal antibody used for inhibiting  
CC tumour growth in an animal. The tumour inhibition process comprises  
CC selecting an animal in need of treatment for a tumour, providing a  
CC monoclonal antibody comprising a heavy chain amino acid, where the  
CC antibody consists of any one of 10 fully defined sequences of 117-123  
CC amino acids given in the specification, and where the monoclonal antibody  
CC binds MUC18, and contacting the tumour with the antibody resulting in  
CC inhibited proliferation of the cells. The monoclonal antibody has  
CC cytostatic and can be used in the production of a vaccine. The monoclonal  
CC antibodies against the MUC18 antigen are useful for diagnosing and  
CC treating tumours, inhibiting tumour growth (e.g. melanoma, lung tumour or  
CC tumour metastasis), inhibiting cell invasion associated with melanoma, or  
CC increasing survival of an animal having a metastatic tumour. This  
CC sequence represents an anti-MUC18 antibody light chain, variable region,  
CC protein of the invention.  
XX  
SQ Sequence 95 AA;  
XX  
Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
XX  
QY 387 ELELTOSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVPA 446  
Db 1 EIVMTQSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVPA 60  
XX  
QY 447 RFGSGSGAEFTLTITSSLSQSEDPFAVYCCQYNNWP 481  
Db 61 RFGSGSGTEFTLTITSSLSQSEDPFAVYCCQYNNWP 95  
XX  
RESULT 8  
ADFI0095  
ID ADFI0095 standard; protein; 95 AA.  
XX  
AC ADFI0095;  
XX  
DT 12-FEB-2004 (first entry)  
XX  
DE VEGF antibody light chain variable region VLK\_3-15.  
XX  
KM Antibody; stability; solubility; antigen binding affinity;  
KM variable region; human; VEGF.  
XX  
OS Homo sapiens.  
XX  
PN WO2003074679-A2.  
XX  
PD 12-SEP-2003.  
XX  
PF 03-MAR-2003; 2003WO-US006598.  
XX  
PR 01-MAR-2002; 2002US-0360843P.  
PR 29-MAY-2002; 2002US-0384197P.  
XX  
PA (XENC-) XENCOR.  
XX  
PI Lazar GA, Desjarlais JR, Marshall SA, Dahiya B;  
XX  
XX WPI; 2003-722066/68.  
XX  
PT Computer optimization of physicochemical properties of antibodies  
PT comprises analyzing the interactions of amino acids at variable  
PT positions.  
XX  
XX Example 6; Fig 16b; 135pp; English.  
XX  
XX The present invention relates to a method for optimizing at least one

CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

CC Sequence 95 AA:

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTOSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLSVSPGERATLSCRAQSQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
 DB 61 RFGSGSGGTEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 9  
 ADF10197  
 ID ADF10197 standard; protein; 95 AA.

AC ADF10197;

DT 12-FEB-2004 (first entry)

DE Antibody light chain variable region VLK\_3-15.

KM Antibody; stability; solubility; antigen binding affinity;

KW variable region; human.

OS Homo sapiens.

PN WO2003074679-A2.

PD 12-SEP-2003.

PF 03-MAR-2003; 2003WO-US006598.

PR 01-MAR-2002; 2002US-0360843P.

PR 29-MAY-2002; 2002US-0384197P.

PA (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

DR WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.

PS Example 16; Fig 40b; 135pp; English.

XX The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;

CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

CC Sequence 95 AA:

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTOSPATLSVSPGERATLSCRASESVSDDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLSVSPGERATLSCRAQSQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
 DB 61 RFGSGSGGTEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 10  
 ADF10198  
 ID ADF10198 standard; protein; 95 AA.

AC ADF10198;

DT 12-FEB-2004 (first entry)

DE Antibody light chain variable region VLK\_3D-15.

KM Antibody; stability; solubility; antigen binding affinity;

KW variable region; human.

OS Homo sapiens.

PN WO2003074679-A2.

PD 12-SEP-2003.

PF 03-MAR-2003; 2003WO-US006598.

PR 01-MAR-2002; 2002US-0360843P.

PR 29-MAY-2002; 2002US-0384197P.

PA (XENC-) XENCOR.

PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;

DR WPI; 2003-722066/68.

PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.

PS Example 16; Fig 40b; 135pp; English.

XX The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.

Seq	Sequence	95 AA:
Query Match	34.0%;	Score 88; DB 1; Length 95;
Best Local Similarity	92.6%;	Pred. No. 6.6;
Matches	88;	Conservative 0; Mismatches 7; Indels 0; Gaps 0;
QY	387	ELELTQSPATISVSGGERATISCRASESVSSPLAWYQOKPGQAPRLLIYGASTRATGVPA 446
DB	1	EIVMTQSPATISVSGGERATISCRASGVSSNLAWYQOKPGQAPRLLIYGASTRATGIPA 60
QY	447	RFGSGSGGAEFTLTITSSLOSEDFAVYYCOQYNWNP 481
DB	61	RFGSGSGCTEFTLTITSSLOSEDFAVYYCOQYNWNP 95
RESULT 11		
ID	ADF09990	standard; protein; 95 AA.
XX	ADF09990;	
XX	ADF09990;	
XX	12-FEB-2004	(first entry)
XX	Antibody light chain variable region VLK_3-15.	
XX	Antibody; stability; solubility; antigen binding affinity;	
XX	variable region; human.	
XX	Homo sapiens.	
XX	WO2003074679-A2.	
XX	12-SEP-2003.	
XX	03-MAR-2003; 2003WO-US0069598.	
XX	01-MAR-2002; 2002US-0360843P.	
XX	29-MAY-2002; 2002US-0384197P.	
XX	(XENC-) XENCOR.	
XX	Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;	
XX	WPI; 2003-722066/68.	
XX	Computer optimization of physicochemical properties of antibodies	
XX	comprises analyzing the interactions of amino acids at variable	
XX	positions.	
XX	Disclosure; Fig 2b; 135p; English.	
XX	The present invention relates to a method for optimizing at least one	
XX	physico-chemical property of an antibody by a computational screening	
XX	method. The method comprises: receiving a template antibody structure;	
XX	selecting at least one variable position belonging to the antibody	
XX	structure; selecting at least one amino acid to be considered at the	
XX	variable position(s); analyzing the interaction of each selected amino	
XX	acid at each variable position with at least part of the remainder of the	
XX	antibody, including the selected amino acids at other variable positions;	
XX	and identifying a set of at least one antibody sequence with at least one	
XX	optimized physico-chemical property. The method is useful for optimizing	
XX	the physico-chemical properties of an antibody, especially the stability,	
XX	solubility, or antigen binding affinity. The optimized antibody may be	
XX	useful for treating a patient. The present sequence is an antibody	
XX	variable region sequence used to illustrate the invention.	
XX	Sequence 95 AA;	
XX	Query Match	34.0%; Score 88; DB 1; Length 95;
XX	Best Local Similarity	92.6%; Pred. No. 6.6;
XX	Matches	88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX	387	ELELTQSPATISVSGGERATISCRASESVSSPLAWYQOKPGQAPRLLIYGASTRATGVPA 446

Db	1	EIVNTQSPATISVSPGERRATISCSASQSVSSNLWYQOKPQADRLIYGASTRATGIPA	60
Qy	447	RFGSGSGAEFTLTISLQSEDFAVYYCQOYNMP	481
Db	61	RFGSGSGTEFTLTISLQSEDFAVYYCQOYNMP	95
RESULT 12			
ID	ADF09991	standard; protein; 95 AA.	
XX	ADF09991;		
AC	ADF09991;		
XX			
XX	12-FEB-2004	(first entry)	
DT			
DE	Antibody light chain variable region VLK_3D-15.		
XX			
XX	Antibody; stability; solubility; antigen binding affinity;		
KW	variable region; human.		
XX			
OS	Homo sapiens.		
XX			
XX	WC0003074679-A2.		
PN			
XX	12-SEP-2003.		
PD			
XX	03-MAR-2003; 2003WC-US006598.		
XX			
FP	01-MAR-2002; 2002US-0360843P.		
PR	29-MAY-2002; 2002US-0384197P.		
XX			
XX	(XENC-) XENCOR.		
PA			
XX			
XX	Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;		
XX			
XX	WPI; 2003-722066/68.		
DR			
XX			
PT	Computer optimization of physicochemical properties of antibodies		
PT	comprises analyzing the interactions of amino acids at variable		
FT	positions.		
XX			
XX			
PS	Disclosure; Fig 2b; 135p; English.		
XX			
XX	The present invention relates to a method for optimizing at least one		
XX	physico-chemical property of an antibody by a computational screening		
CC	method. The method comprises: receiving an antibody structure;		
CC	selecting at least one variable position belonging to the antibody		
CC	structure; selecting at least one amino acid to be considered at the		
CC	variable position(s); analyzing the interaction of each selected amino		
CC	acid at each variable position with at least part of the remainder of the		
CC	antibody, including the selected amino acids at other variable positions;		
CC	and identifying a set of at least one antibody sequence with at least one		
CC	optimized physico-chemical property. The method is useful for optimizing		
CC	the physico-chemical properties of an antibody, especially for stabilizing		
CC	solubility, or antigen binding affinity. The optimized antibody may be		
CC	useful for treating a patient. The present sequence is an antibody		
CC	variable region sequence used to illustrate the invention.		
XX			
XX	Sequence 95 AA;		
XX			
XX	Query Match	34.0%; Score 88; DB 1; Length 95;	
XX	Best Local Similarity	92.6%; Pred. No. 6.6;	
XX	Matches	88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;	
Qy	387	ELELTQSPATISVSPGERRATISCSASQSVSSDLAWYQOKPQADRLIYGASTRATGIPA	446
Db	1	EIVNTQSPATISVSPGERRATISCSASQSVSSNLWYQOKPQADRLIYGASTRATGIPA	60
Qy	447	RFGSGSGAEFTLTISLQSEDFAVYYCQOYNMP	481
Db	61	RFGSGSGTEFTLTISLQSEDFAVYYCQOYNMP	95

RESULT 13  
 ADF10096  
 ID ADF10096 standard; protein; 95 AA.  
 AC ADF10096;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE VEGF antibody light chain variable region VLK\_3D-15.  
 DE  
 XX  
 KW Antibody; stability; solubility; antigen binding affinity;  
 KW variable region; human; VEGF.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003074679-A2.  
 PD  
 XX 12-SEP-2003.  
 XX  
 PF 03-MAR-2003; 2003WO-US006598.  
 XX  
 PR 01-MAR-2002; 2002US-0360843P.  
 PR 29-MAY-2002; 2002US-0384197P.  
 XX  
 PA (XENC-) XENCOR.  
 XX  
 PI Lazar GA, Desjarlais JR, Marshall SA, Dahiyat B;  
 XX  
 DR WPI; 2003-722066/68.  
 XX  
 PT Computer optimization of physicochemical properties of antibodies  
 PT comprises analyzing the interactions of amino acids at variable  
 PT positions.  
 XX  
 PS Example 6; Fig 16b; 135pp; English.  
 XX  
 CC The present invention relates to a method for optimizing at least one  
 CC physico-chemical property of an antibody by a computational screening  
 CC method. The method comprises: receiving a template antibody structure;  
 CC selecting at least one variable position belonging to the antibody  
 CC structure; selecting at least one amino acid to be considered at the  
 CC variable position(s); analyzing the interaction of each selected amino  
 CC acid at each variable position with at least part of the remainder of the  
 CC antibody, including the selected amino acids at other variable positions;  
 CC and identifying a set of at least one antibody sequence with at least one  
 CC optimized physico-chemical property. The method is useful for optimizing  
 CC the physico-chemical properties of an antibody, especially the stability,  
 CC solubility, or antigen binding affinity. The optimized antibody may be  
 CC useful for treating a patient. The present sequence is an antibody  
 CC variable region sequence used to illustrate the invention.  
 XX  
 SQ Sequence 95 AA:  
 XX  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTGSPATLSVSPGERATLSCRASSVSSDLAWYQQRGQAPRLIIYGASTRATGVP 446  
 DB 1 EIVMTGSPATLSVSPGERATLSCRASSVSSDLAWYQQRGQAPRLIIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTISLSQSEDPAYVYCCOQYNWNP 481  
 DB 61 RFSGSGSGTEFTLTISLSQSEDPAYVYCCOQYNWNP 95  
 QY 447 RFSGSGSGAEFTLTISLSQSEDPAYVYCCOQYNWNP 481  
 DB 61 RFSGSGSGTEFTLTISLSQSEDPAYVYCCOQYNWNP 95  
 RESULT 14  
 ADF09880  
 ID ADF09880 standard; protein; 95 AA.  
 AC ADF09880;  
 XX

DT 12-FEB-2004 (first entry)  
 XX  
 DE Anti-MUC18 monoclonal antibody-related protein #14.  
 DE  
 XX  
 KW cell proliferation inhibition; MUC18 tumour antigen;  
 KW anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;  
 KW carcinoma; cancer; malignancy.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003057837-A2.  
 PD  
 XX 17-JUL-2003.  
 XX  
 PF 26-DEC-2002; 2002WO-US041580.  
 XX  
 PR 28-DEC-2001; 2001US-0346414P.  
 XX  
 PA (ABGE-) ABGENIX INC.  
 XX  
 PI Gudae J;  
 XX  
 DR WPI; 2003-598367/56.  
 XX  
 PT Inhibiting cell proliferation associated with expression of MUC18 tumor  
 PT antigen, involves incubating and inhibiting cell by administering anti-  
 PT MUC18 monoclonal antibody.  
 XX  
 PS Example 2; SEQ ID NO 67; 83pp; English.  
 XX  
 CC The invention comprises a method for inhibiting cell proliferation  
 CC associated with expression of MUC18 tumour antigen. The method involves  
 CC administering anti-MUC18 monoclonal antibody. The method of the invention  
 CC is useful for inhibiting cell (e.g. melanoma or tumour cell)  
 CC proliferation associated with the expression of MUC18 tumour antigen, the  
 CC method is preferably useful for inhibiting tumour metastasis. The method  
 CC is useful for inhibiting cell proliferation in patients with tumours,  
 CC carcinomas, cancer and other malignancies. The present amino acid  
 CC sequence is used in an alignment with an MUC18 tumour antigen-specific  
 CC monoclonal antibody of the invention.  
 XX  
 SQ Sequence 95 AA:  
 XX  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 ELELTGSPATLSVSPGERATLSCRASSVSSDLAWYQQRGQAPRLIIYGASTRATGVP 446  
 DB 1 EIVMTGSPATLSVSPGERATLSCRASSVSSDLAWYQQRGQAPRLIIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTISLSQSEDPAYVYCCOQYNWNP 481  
 DB 61 RFSGSGSGTEFTLTISLSQSEDPAYVYCCOQYNWNP 95  
 QY 447 RFSGSGSGAEFTLTISLSQSEDPAYVYCCOQYNWNP 481  
 DB 61 RFSGSGSGTEFTLTISLSQSEDPAYVYCCOQYNWNP 95  
 RESULT 15  
 ADF09888  
 ID ADF09888 standard; protein; 95 AA.  
 AC ADF09888;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Anti-MUC18 monoclonal antibody-related protein #18.  
 DE  
 XX  
 KW cell proliferation inhibition; MUC18 tumour antigen;  
 KW anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;  
 KW carcinoma; cancer; malignancy.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003057837-A2.

XX 17-JUL-2003.  
 PD 26-DEC-2002; 2002WO-US041580.  
 XX 28-DEC-2001; 2001US-0346414P.  
 PR (ABGE-) ABGENIX INC.  
 XX  
 PA Gudae J;  
 PI WPI; 2003-598367/56.  
 DR Inhibiting cell proliferation associated with expression of MUC18 tumor  
 XX antigen, involves incubating and inhibiting cell by administering anti-  
 PT MUC18 monoclonal antibody.  
 CC Example 2; SEQ ID NO 75; 83pp; English.  
 PS  
 XX The invention comprises a method for inhibiting cell proliferation  
 CC associated with expression of MUC18 tumour antigen. The method involves  
 CC administering anti-MUC18 monoclonal antibody. The method of the invention  
 CC is useful for inhibiting cell (e.g. melanoma or tumour cell)  
 CC proliferation associated with the expression of MUC18 tumour antigen, the  
 CC method is preferably useful for inhibiting tumour metastasis. The method  
 CC is useful for inhibiting cell proliferation in patients with tumours,  
 CC carcinomas, cancer and other malignancies. The present amino acid  
 CC sequence is used in an alignment with an MUC18 tumour antigen-specific  
 CC monoclonal antibody of the invention.  
 XX  
 SQ Sequence 95 AA;  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 EELTQSPATLVSFGERATLSCRASVSVDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLVSFGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWP 481  
 DB 61 RFSGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP 95  
 RESULT 16  
 ADF09892  
 ID ADF09892 standard; protein; 95 AA.  
 XX  
 AC ADF09892;  
 XX  
 DT 12-FEB-2004 (first entry)  
 XX  
 DE Anti-MUC18 monoclonal antibody-related protein #20.  
 XX  
 KW cell proliferation inhibition; MUC18 tumour antigen;  
 KM anti-MUC18 monoclonal antibody; tumour metastasis inhibition; tumour;  
 KM carcinoma; cancer; malignancy.  
 XX  
 OS Unidentified.  
 XX  
 PN WO2003057837-A2.  
 PD 17-JUL-2003.  
 XX  
 PF 26-DEC-2002; 2002WO-US041580.  
 XX  
 PR 28-DEC-2001; 2001US-0346414P.  
 XX  
 PA (ABGE-) ABGENIX INC.  
 XX  
 PI Gudae J;  
 XX

DR WPI; 2003-598367/56.  
 XX Inhibiting cell proliferation associated with expression of MUC18 tumor  
 PT antigen, involves incubating and inhibiting cell by administering anti-  
 PT MUC18 monoclonal antibody.  
 CC Example 2; SEQ ID NO 79; 83pp; English.  
 PS  
 XX The invention comprises a method for inhibiting cell proliferation  
 CC associated with expression of MUC18 tumour antigen. The method involves  
 CC administering anti-MUC18 monoclonal antibody. The method of the invention  
 CC is useful for inhibiting cell (e.g. melanoma or tumour cell)  
 CC proliferation associated with the expression of MUC18 tumour antigen, the  
 CC method is preferably useful for inhibiting tumour metastasis. The method  
 CC is useful for inhibiting cell proliferation in patients with tumours,  
 CC carcinomas, cancer and other malignancies. The present amino acid  
 CC sequence is used in an alignment with an MUC18 tumour antigen-specific  
 CC monoclonal antibody of the invention.  
 XX  
 SQ Sequence 95 AA;  
 Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
 QY 387 EELTQSPATLVSFGERATLSCRASVSVDLAWYQKPGQAPRLIYGASTRATGVA 446  
 DB 1 EIVMTQSPATLVSFGERATLSCRASQSVSNLAWYQKPGQAPRLIYGASTRATGIPA 60  
 QY 447 RFSGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWP 481  
 DB 61 RFSGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP 95  
 RESULT 17  
 ADJ80271  
 ID ADJ80271 standard; protein; 95 AA.  
 XX  
 AC ADJ80271;  
 XX  
 DT 06-MAY-2004 (first entry)  
 XX  
 DE V kappa gene locus amino acid sequence #31.  
 XX  
 KW hybrid antibody; antibody; framework region; homology; immunogenicity.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO2003048321-A2.  
 PD 12-JUN-2003.  
 XX  
 PF 03-DEC-2002; 2002WO-US038450.  
 XX  
 PR 03-DEC-2001; 2001US-0336591P.  
 XX  
 PA (ALEX-) ALEXION PHARM INC.  
 XX  
 PI Rother R, Wu D;  
 XX  
 DR WPI; 2003-513753/48.  
 XX  
 PT Producing a hybrid antibody or hybrid antibody fragment by operatively  
 PT linking the selected framework sequences to one or more complementarity  
 PT determining regions of the initial antibody.  
 XX  
 PS Disclosure; SEQ ID NO 31; 77pp; English.  
 XX  
 CC The invention relates to a method of producing a hybrid antibody or  
 CC hybrid antibody fragment by: (i) providing an initial antibody having  
 CC specificity for a target; (ii) determining the sequence of a variable  
 CC region of the initial antibody; (iii) selecting a first component of the  
 CC variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the

CC sequence of the first component to sequences contained in a reference  
CC database of antibody sequences or antibody fragment sequences from a  
CC target species; (v) selecting a sequence from an antibody in the database  
CC which demonstrates a high degree of homology to the first component; (vi)  
CC selecting a second component of the variable region which is different  
CC than the first component, the second component selected from the group  
CC consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the  
CC second component to sequences contained in a reference database of  
CC antibody sequences or antibody fragment sequences from the target species  
CC ; (viii) selecting a sequence from the database which demonstrates a high  
CC degree of homology to the second component and which is from a different  
CC antibody than the selected antibody; and (ix) operatively linking the  
CC selected framework sequences to one or more complementarily determining  
CC regions (CDRs) of the initial antibody to produce a hybrid antibody or  
CC hybrid antibody fragment. The method is useful for producing a hybrid  
CC antibody or hybrid antibody fragment (claimed). The antibody and  
CC fragments are useful for therapeutic and diagnostic purposes. The method  
CC uses entire framework regions from a single antibody variable heavy or  
CC variable light chain to receive the CDRs. This produces antibodies that  
CC are highly homologous and exhibit reduced immunogenicity while  
CC maintaining an optimum binding profile. This sequence represents the  
CC amino acid sequence of an antibody from the V kappa gene locus.

XX Sequence 95 AA;

SQ Query Match 34.0%; Score 88; DB 1; Length 95;

Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTGSPATLSVSGERATLSCRASBSVSSDLAWYQKQKQAPRLIYGASTRATGVA 446

Db 1 EIWVTGSPATLSVSGERATLSCRASQSVSSNLAWYQKQKQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNNMP 481

Db 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNNMP 95

## RESULT 18

ADJ80272 standard; protein; 95 AA.

XX AC ADJ80272;

DT 06-MAY-2004 (first entry)

DE V kappa gene locus antibody amino acid sequence #32.

XX hybrid antibody; antibody; framework region; homology; immunogenicity.

XX Homo sapiens.

XX PN WO2003048321-A2.

XX PD 12-JUN-2003.

PF 03-DEC-2002; 2002MO-US038450.

PR 03-DEC-2001; 2001US-0336591P.

PA (ALEX-) ALEXION PHARM INC.

PI Rother R, Wu D;

DR WPI; 2003-513753/48.

XX Producing a hybrid antibody or hybrid antibody fragment by operatively  
PT linking the selected framework sequences to one or more complementarily  
PT determining regions of the initial antibody.

XX Disclosure; SEQ ID NO 32; 77pp; English.

XX The invention relates to a method of producing a hybrid antibody or

CC hybrid antibody fragment by: (i) providing an initial antibody having  
CC specificity for a target; (ii) determining the sequence of a variable  
CC region of the initial antibody; (iii) selecting a first component of the  
CC variable region consisting of FR1, FR2, FR3 and FR4; (iv) comparing the  
CC sequence of the first component to sequences contained in a reference  
CC database of antibody sequences or antibody fragment sequences from a  
CC target species; (v) selecting a sequence from an antibody in the database  
CC which demonstrates a high degree of homology to the first component; (vi)  
CC selecting a second component of the variable region which is different  
CC than the first component, the second component selected from the group  
CC consisting of FR1, FR2, FR3 and FR4; (vii) comparing the sequence of the  
CC second component to sequences contained in a reference database of  
CC antibody sequences or antibody fragment sequences from the target species  
CC ; (viii) selecting a sequence from the database which demonstrates a high  
CC degree of homology to the second component and which is from a different  
CC antibody than the selected antibody; and (ix) operatively linking the  
CC selected framework sequences to one or more complementarily determining  
CC regions (CDRs) of the initial antibody to produce a hybrid antibody or  
CC hybrid antibody fragment. The method is useful for producing a hybrid  
CC antibody or hybrid antibody fragment (claimed). The antibody and  
CC fragments are useful for therapeutic and diagnostic purposes. The method  
CC uses entire framework regions from a single antibody variable heavy or  
CC variable light chain to receive the CDRs. This produces antibodies that  
CC are highly homologous and exhibit reduced immunogenicity while  
CC maintaining an optimum binding profile. This sequence represents the  
CC amino acid sequence of an antibody from the V kappa gene locus.

SQ Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;

Best Local Similarity 92.6%; Pred. No. 6.6;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELELTGSPATLSVSGERATLSCRASBSVSSDLAWYQKQKQAPRLIYGASTRATGVA 446

Db 1 EIWVTGSPATLSVSGERATLSCRASQSVSSNLAWYQKQKQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNNMP 481

Db 61 RFGSGSGTEFTLTISLSQSEDFAVYYCCQYNNMP 95

## RESULT 19

ADJ007341 standard; protein; 95 AA.

XX AC ADJ007341;

DT 15-JUL-2004 (first entry)

DE Human antibody L16 light chain variable region.

XX Catalytic antibody; human; antibody; gene therapy.

XX Homo sapiens.

XX Location/Qualifiers

FT 1..23 /label= FR1

FT 24..34 /note= "Framework region 1"

FT /label= CDR1 /note= "Complementarity determining region 1"

FT 35..49 /label= FR2 /note= "Framework region 2"

FT 50..56 /label= CDR2 /note= "Complementarity determining region 2"

FT 57..68 /label= FR3

FT 69..89 /note= "Framework region 3"

FT Region

```

FT      /label= CDR3
XX      /note= "Complementarity determining region 3"
XX      MO2004033658-A2.
XX
XX      22-APR-2004.
XX
XX      09-OCT-2003; 2003WO-US032214.
XX
XX      10-OCT-2002; 2002US-0417979P.
XX
XX      (INTE-) INTEGRIGEN INC.
XX
XX      Smider V, Larrick JW;
XX
XX      WPI; 2004-340921/31.
XX
XX      Novel recombinant catalytic polypeptide useful for cleaving target
PT      proteins or for treating or preventing cancers, comprises a human
PT      antibody light chain operably joined to a heterologous antibody heavy
PT      chain.
XX
XX      Disclosure; Fig 3; 92pp; English.
XX
XX      The present sequence is the light chain variable region of human antibody
CC      116. This is one of repertoire of human kappa light chain sequences
CC      ADO07310-ADO07349 screened for putative catalytic triads. Several genes
CC      encoding such light chains ADO07282-ADO07309 were cloned for use in
CC      recombinant catalytic polypeptides of the invention. These comprise a
CC      human antibody light chain operably joined to a heterologous antibody
CC      heavy chain. The light chain has a predetermined specificity for a
CC      target protein. By joining 2 heterologous human antibody chains, one of
CC      which supplies the catalytic activity to hydrolyse polypeptides and the
CC      other the binding specificity for a target protein, the invention
CC      provides for the construction of a repertoire of proteases with
CC      customised protein substrate specificities of potentially unlimited
CC      number and thus makes possible the effective treatment and/or prevention
CC      of any medical condition attributable to the presence or overexpression
CC      of an identified protein. The invention also provides nucleic acids
CC      encoding the catalytic antibodies (which can be used for gene therapy),
CC      host cells, transgenic non-human animals, and methods of cleaving a
CC      target protein (in vitro or in vivo) using a recombinant catalytic
CC      polypeptide. It also provides a library of recombinant catalytic
CC      polypeptides with altered enzymatic activity, and a method of altering
CC      the enzymatic activity of the recombinant catalytic polypeptides by
CC      mutating at least one complementarity determining region of the heavy
CC      chain. A serine protease triad was not identified in the present
CC      sequence.
XX
XX      Sequence 95 AA;
XX
XX      Query Match      34.0%; Score 88; DB 1; Length 95;
XX      Best Local Similarity 92.6%; Pred. No. 6.6;
XX      Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;
XX
XX      387  ELELTGSPATLVSFGERATLSCRASSVSVDLAWYQKQKQAPRLIYCASTRATGVA 446
XX      1  EIVWTGSPATLVSFGERATLSCRASSVSVDLAWYQKQKQAPRLIYCASTRATGIP 60
XX
XX      447  RFGSGSGAFTLTISLQSEDFAVYYCCQYNMP 481
XX      61  RFGSGSGAFTLTISLQSEDFAVYYCCQYNMP 95
XX
XX      RESULT 20
XX      ID ADO07340 standard; protein; 95 AA.
XX
XX      ADO07340;
XX
XX      15-JUL-2004 (first entry)
XX

```

```

DE      Human antibody L2 light chain variable region.
XX
XX      Catalytic antibody; human; antibody; gene therapy.
XX
XX      Homo sapiens.
XX
XX      Key
XX      Location/Qualifiers
XX      Region
XX      1..23
XX      /label= FR1
XX      /note= "Framework region 1"
XX
XX      Region
XX      24..34
XX      /label= CDR1
XX      /note= "Complementarity determining region 1"
XX
XX      Region
XX      35..49
XX      /label= FR2
XX      /note= "Framework region 2"
XX
XX      Region
XX      50..56
XX      /label= CDR2
XX      /note= "Complementarity determining region 2"
XX
XX      Region
XX      57..88
XX      /label= FR3
XX      /note= "Framework region 3"
XX
XX      Region
XX      89..95
XX      /label= CDR3
XX      /note= "Complementarity determining region 3"
XX
XX      MO2004033658-A2.
XX
XX      22-APR-2004.
XX
XX      09-OCT-2003; 2003WO-US032214.
XX
XX      10-OCT-2002; 2002US-0417979P.
XX
XX      (INTE-) INTEGRIGEN INC.
XX
XX      Smider V, Larrick JW;
XX
XX      WPI; 2004-340921/31.
XX
XX      Novel recombinant catalytic polypeptide useful for cleaving target
PT      proteins or for treating or preventing cancers, comprises a human
PT      antibody light chain operably joined to a heterologous antibody heavy
PT      chain.
XX
XX      Disclosure; Fig 3; 92pp; English.
XX
XX      The present sequence is the light chain variable region of human antibody
CC      116. This is one of repertoire of human kappa light chain sequences
CC      ADO07310-ADO07349 screened for putative catalytic triads. Several genes
CC      encoding such light chains ADO07282-ADO07309 were cloned for use in
CC      recombinant catalytic polypeptides of the invention. These comprise a
CC      human antibody light chain operably joined to a heterologous antibody
CC      heavy chain. The light chain has a predetermined specificity for a
CC      target protein. By joining 2 heterologous human antibody chains, one of
CC      which supplies the catalytic activity to hydrolyse polypeptides and the
CC      other the binding specificity for a target protein, the invention
CC      provides for the construction of a repertoire of proteases with
CC      customised protein substrate specificities of potentially unlimited
CC      number and thus makes possible the effective treatment and/or prevention
CC      of any medical condition attributable to the presence or overexpression
CC      of an identified protein. The invention also provides nucleic acids
CC      encoding the catalytic antibodies (which can be used for gene therapy),
CC      host cells, transgenic non-human animals, and methods of cleaving a
CC      target protein (in vitro or in vivo) using a recombinant catalytic
CC      polypeptide. It also provides a library of recombinant catalytic
CC      polypeptides with altered enzymatic activity, and a method of altering
CC      the enzymatic activity of the recombinant catalytic polypeptides by
CC      mutating at least one complementarity determining region of the heavy
CC      chain. A serine protease triad was not identified in the present
CC      sequence.
XX

```

SQ Sequence 95 AA;

Query Match 34.0%; Score 88; DB 1; Length 95;  
 Best Local Similarity 92.6%; Pred. No. 6.6;  
 Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

OY 387 ELELTGSPATLSVSPGERATLSCPASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVPA 446  
 DB 1 EIVWTGSPATLSVSPGERATLSCPASBSVSSDLAWYQKPGQAPRLLIYGASTRATGVPA 60  
 OY 447 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNMP 481  
 DB 61 RFGSGSGAEFTLTISLSQSEDFAVYYCCQYNMP 95

RESULT 21

ID ADR43440 standard; peptide: 32 AA.

AC ADR43440;

DT 04-NOV-2004 (first entry)

DE FRL variant #11.

XX antibody; variable light chain; variable heavy chain; Antiallergic;  
 KM Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
 KM eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.

OS Unidentified.

PN WO2004070011-A2.

PD 19-AUG-2004.

PF 02-FEB-2004; 2004WO-US002894.

PR 01-FEB-2003; 2003US-0444229P.

PA (TANO-) TANOX INC.

PI Singh S, Foster C, Wu H;

PP WPI; 2004-604433/58.

PT New high affinity human monoclonal antibodies, particularly those  
 directed against isotypic determinants of immunoglobulin E, useful for  
 asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
 food allergy.

PS Claim 1; SEQ ID NO 40; 101pp; English.

CC The present invention relates to an antibody comprising a variable light  
 chain region or a variable heavy chain region. The antibody and methods  
 are useful for treating a disorder associated with an abnormally high IGE  
 level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
 dermatitis, or a food allergy. The present sequence represents a FRL  
 which demonstrates improved affinity. This sequence is present in the  
 library of variants in the specification.

SQ Sequence 32 AA;

Query Match 11.6%; Score 30; DB 1; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 27;  
 Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 443 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 474  
 DB 1 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 32

RESULT 22

ADR31585

ID ADR31585 standard; peptide: 32 AA.

AC ADR31585;

DT 04-NOV-2004 (first entry)

DE Murine TES-C21 antibody (clone 136) FRK3 variant peptide.

KM Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
 VK; murine; FR; framework region; mutant; mutein; variant.

OS Mus sp.

PN WO2004070010-A2.

PD 19-AUG-2004.

PF 02-FEB-2004; 2004WO-US002892.

PR 01-FEB-2003; 2003US-0444229P.

PA (TANO-) TANOX INC.

PI Singh S, Foster C, Wu H;

PP WPI; 2004-604432/58.

PT Generating a humanized, high affinity antibody from an antibody of  
 PT interest comprises selecting a suitable human template as the framework  
 for the H and L chain variable domains of the high affinity antibody to  
 be made.

PS Example 10; SEQ ID NO 40; 100pp; English.

CC The invention relates to a method for generating a humanised high  
 CC affinity antibody from an antibody of interest. The method involves  
 CC selecting a suitable human template as the framework for the H (heavy)  
 CC and L (light) chain variable (V) domains of the high affinity antibody to  
 CC be made. The method is useful for generating high affinity antibodies  
 CC useful in diagnostics, prophylaxis and treatment of diseases. The present  
 CC sequence is murine anti-human IGE antibody (TES-C21) VK (kappa chain  
 CC variable region) FR (framework region) variant peptide.

SQ Sequence 32 AA;

Query Match 11.6%; Score 30; DB 1; Length 32;  
 Best Local Similarity 93.8%; Pred. No. 27;  
 Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 443 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 474  
 DB 1 GVPARFSGSGAEFTLTISLSQSEDFAVYYC 32

RESULT 23

ID ADR43442 standard; peptide: 32 AA.

AC ADR43442;

DT 04-NOV-2004 (first entry)

DE FRL variant #13.

XX antibody; variable light chain; variable heavy chain; Antiallergic;  
 KM Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
 KM eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.

OS Unidentified.

PN WO2004070011-A2.

PD 19-AUG-2004.  
XX  
XX 02-FEB-2004; 2004WO-US002894.  
XX  
XX 01-FEB-2003; 2003US-0444229P.  
XX  
XX (TANQ-) TANOX INC.  
XX  
XX Singh S, Foster C, Wu H;  
XX  
XX WPI; 2004-604433/58.  
XX  
XX  
XX New high affinity human monoclonal antibodies, particularly those  
XX directed against isotypic determinants of immunoglobulin E, useful for  
XX asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
XX food allergy.  
XX  
XX Claim 1; SEQ ID NO 42; 101pp; English.  
XX  
XX The present invention relates to an antibody comprising a variable light  
XX chain region or a variable heavy chain region. The antibody and methods  
XX are useful for treating a disorder associated with an abnormally high IGE  
XX level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
XX dermatitis, or a food allergy. The present sequence represents a FRL  
XX which demonstrates improved affinity. This sequence is present in the  
XX library of variants in the specification.  
XX  
XX Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
Best Local Similarity 90.6%; Pred. No. 27;  
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 443 GVPARFSGSGGAEFTLTSSLOSEDPAYYC 474  
Db 1 GIPSRFSGSGGTEFTLTSSLOSEDPADYYC 32

RESULT 24  
ADR43441  
ID ADR43441 standard; peptide; 32 AA.  
XX  
XX ADR43441;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX FRL variant #12.  
XX  
XX antibody; variable light chain; variable heavy chain; Anti-allergic;  
XX Dermatological; Immunosuppressive; IGE; asthma; allergic rhinitis;  
XX eczema; urticaria; atopic dermatitis; food allergy; CDR; FRL.  
XX  
XX Undifferentiated.  
XX  
XX WO2004070011-A2.  
XX  
XX 19-AUG-2004.  
XX  
XX 02-FEB-2004; 2004WO-US002894.  
XX  
XX 01-FEB-2003; 2003US-0444229P.  
XX  
XX (TANQ-) TANOX INC.  
XX  
XX Singh S, Foster C, Wu H;  
XX  
XX WPI; 2004-604433/58.  
XX  
XX  
XX New high affinity human monoclonal antibodies, particularly those  
XX directed against isotypic determinants of immunoglobulin E, useful for  
XX asthma, allergic rhinitis, eczema, urticaria, atopic dermatitis, or a  
XX food allergy.

PS Claim 1; SEQ ID NO 41; 101pp; English.  
XX  
XX The present invention relates to an antibody comprising a variable light  
XX chain region or a variable heavy chain region. The antibody and methods  
XX are useful for treating a disorder associated with an abnormally high IGE  
XX level, e.g. asthma, allergic rhinitis, eczema, urticaria, atopic  
XX dermatitis, or a food allergy. The present sequence represents a FRL  
XX which demonstrates improved affinity. This sequence is present in the  
XX library of variants in the specification.  
XX  
XX Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
Best Local Similarity 90.6%; Pred. No. 27;  
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
Qy 443 GVPARFSGSGGAEFTLTSSLOSEDPAYYC 474  
Db 1 GIPSRFSGSGGTEFTLTSSLOSEDPAYYC 32

RESULT 25  
ADR31587  
ID ADR31587 standard; peptide; 32 AA.  
XX  
XX ADR31587;  
XX  
XX 04-NOV-2004 (first entry)  
XX  
XX Murine TES-C21 antibody (clone 13) FRK3 variant peptide.  
XX  
XX Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
XX VK; murine; FR; framework region; mutant; mutein; variant.  
XX  
XX Mus sp.  
XX Synthetic.  
XX  
XX WO2004070010-A2.  
XX  
XX 19-AUG-2004.  
XX  
XX 02-FEB-2004; 2004WO-US002892.  
XX  
XX 01-FEB-2003; 2003US-0444229P.  
XX  
XX (TANQ-) TANOX INC.  
XX  
XX Singh S, Foster C, Wu H;  
XX  
XX WPI; 2004-604432/58.  
XX  
XX  
XX Generating a humanized, high affinity antibody from an antibody of  
XX interest comprises selecting a suitable human template as the framework  
XX for the H and L chain variable domains of the high affinity antibody to  
XX be made.  
XX  
XX Example 10; SEQ ID NO 42; 100pp; English.  
XX  
XX The invention relates to a method for generating a humanized high  
XX affinity antibody from an antibody of interest. The method involves  
XX selecting a suitable human template as the framework for the H (heavy)  
XX and L (light) chain variable (V) domains of the high affinity antibody to  
XX be made. The method is useful for generating high affinity antibodies  
XX useful in diagnostics, prophylaxis and treatment of diseases. The present  
XX sequence is murine anti-human IGE antibody (TES-C21) VK (kappa chain  
XX variable region) FR (framework region) variant peptide.  
XX  
XX Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
Best Local Similarity 90.6%; Pred. No. 27;  
Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGSGAEFTLTISLSLQSEDPAYVYC 474  
 DB 1 GIPARFSGSGSGTEFTLTISLSLQSEDPADYVC 32

RESULT 26  
 ID ADR31586 standard; peptide; 32 AA.

AC ADR31586;

DT 04-NOV-2004 (first entry)

DE Murine TES-C21 antibody (clone 1) FRK3 variant peptide.

KW Antibody; diagnostic; prophylaxis; therapy; kappa chain variable region;  
 KM VK; murine; FR; framework region; mutant; mutein; variant.

OS Mus sp.  
 OS Synthetic.

XX WO2004070010-A2.

XX 19-AUG-2004.

XX 02-FEB-2004; 2004MO-US002892.

XX 01-FEB-2003; 2003US-0444229P.

XX (TANCO-) TANCOX INC.

XX Singh S, Foster C, Wu H;

XX WPI; 2004-604432/58.

XX Generating a humanized, high affinity antibody from an antibody of  
 PT interest comprises selecting a suitable human template as the framework  
 PT for the H and L chain variable domains of the high affinity antibody to  
 be made.

XX Example 10; SEQ ID NO 41; 100pp; English.

XX The invention relates to a method for generating a humanised high  
 CC affinity antibody from an antibody of interest. The method involves  
 CC selecting a suitable human template as the framework for the H (heavy)  
 CC and L (light) chain variable (V) domains of the high affinity antibodies  
 CC to be made. The method is useful for generating high affinity antibodies  
 CC useful in diagnostics, prophylaxis and treatment of diseases. The present  
 CC sequence is murine anti-human Igs antibody (TES-C21) VK (kappa chain  
 CC variable region) FR (framework region) variant peptide.

XX Sequence 32 AA;

Query Match 11.2%; Score 29; DB 1; Length 32;  
 Best Local Similarity 90.6%; Pred. No. 27;

Matches 29; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 443 GVPARFSGSGSGAEFTLTISLSLQSEDPAYVYC 474  
 DB 1 GIPARFSGSGSGTEFTLTISLSLQSEDPAYVYC 32

RESULT 27

ID AAU90894 standard; peptide; 26 AA.

AC AAU90894;

DT 18-JUN-2002 (first entry)

DE Insulin/insulin-like growth factor receptor-binding peptide #2850.

XX Cytostatic; antidiabetic; neuroprotective; cerebroprotective;

KW ophthalmological; insulin; receptor; gene therapy; diabetes;  
 KM insulin-like growth factor-1; IGF-1; tumour; prostate; breast;  
 KM diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.

OS Synthetic.

XX WO200172771-A2.

XX 04-OCT-2001.

XX 29-MAR-2000; 2000MO-US008528.

XX 29-MAR-2000; 2000MO-US008528.

XX (DGI-) DGI BIOTECHNOLOGIES LLC.

XX (NOVO) NOVO NORDISK AS.

XX Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;  
 PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Mandeckl WS;  
 PI Hansen PH, Ravera M, Hsiao K;

XX WPI; 2002-025774/03.

XX Modulating insulin activity in mammalian cells, for treating e.g.  
 PT diabetes and tumors, comprises using peptides that bind to insulin or  
 PT insulin-like growth factor receptors.

XX Example 5; Fig 33A; 390pp; English.

XX The invention relates to a method of modulating insulin activity in  
 CC mammalian cells by administering a peptide that binds the insulin  
 CC receptor (IR). A composition containing a peptide, optionally expressed  
 CC from gene therapy vectors, that binds to site 1 of IR and an insulin  
 CC agonist are useful for treating diabetes. Also, peptides that are  
 CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are  
 CC useful for treating insulin-like growth factor (IGF)-sensitive tumours  
 CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1  
 CC receptor agonists are useful for treating neurological diseases.

XX Including stroke and diabetic neuropathy. The peptides are also useful in  
 CC screening for compounds that bind to IR or IGF-1 receptor, potential  
 CC therapeutics and research reagents. AAU88034-AAU90957 represent IR and/or  
 CC IGF-1 receptor-binding peptides and related amino acid sequences of the  
 CC invention

XX Sequence 26 AA;

Query Match 9.3%; Score 24; DB 1; Length 26;  
 Best Local Similarity 92.3%; Pred. No. 27;

Matches 24; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 361 WGGGTLVTVTSSGGSGSGSGSGSGS 386  
 DB 1 WGGGTLVTVSSGGSGSGSGSGSGS 26

Search completed: September 28, 2005, 10:28:41  
 Job time : 1 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 28, 2005, 10:30:43 ; Search time 0.001 Seconds  
(without alignments)  
22.274 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPGSSVKV.....RYFGQGRLEIKLVPRGSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapext 0.5

Searched: 5 seqs, 86 residues

Total number of hits satisfying chosen parameters: 5

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 5 summaries

Database: rai.subdb.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	18	6.9	19	1	PCT-US91-02942-67
2	17	6.6	18	1	US-08-207-169A-7
3	16	6.2	16	1	US-08-331-398A-54
4	16	6.2	16	1	US-08-331-397B-54
5	16	6.2	17	1	US-08-463-163-5

#### ALIGNMENTS

RESULT 1  
PCT-US91-02942-67  
Sequence 67, Application PC/TUS9102942  
GENERAL INFORMATION:  
APPLICANT: ROTHLEIN, ROBERT  
APPLICANT: ADAIR, JOHN R  
APPLICANT: ATHWAL, DILJEET S  
TITLE OF INVENTION: HUMANIZED CDR-GRAFTED ICAM-1 ANTIBODY  
NUMBER OF SEQUENCES: 102  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Sterne, Kessler, Goldstein & Fox  
STREET: 1225 Connecticut Ave. NW Suite 300  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US91/02942

FILING DATE: 19910429  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9009549.8  
FILING DATE: 27-APR-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: FOX, SAM L  
REGISTRATION NUMBER: 30,353  
REFERENCE/DOCKET NUMBER: 1011.0586600  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 466-0800  
TELEFAX: (202) 833-8716  
INFORMATION FOR SEQ ID NO: 67:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19 amino acids  
TYPE: AMINO ACID  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
PCT-US91-02942-67

Query Match 6.9%; Score 18; DB 1; Length 19;  
Best Local Similarity 100.0%; Pred. No. 0.29;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 250 SGAEVKKPGSSVKSCKA 267  
Db 2 SGAEVKKPGSSVKSCKA 19

RESULT 2  
US-08-207-169A-7  
Sequence 7, Application US/08207169A  
Patent No. 5674712  
GENERAL INFORMATION:  
APPLICANT: GRANDI, GUIDO  
APPLICANT: DE FERRA, FRANCESCA  
APPLICANT: TOSI, CLAUDIO  
APPLICANT: TORTORA, ORNELLA  
APPLICANT: CUZZONI, ANNA  
TITLE OF INVENTION: RECOMBINANT VECTOR AND USE THEREOF FOR  
TITLE OF INVENTION: EXOCELLULAR PREPARATION OF SINGLE MOLECULA ANTIBODIES FROM  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,  
ADDRESSER: P.C.  
STREET: 1755 S. Jefferson Davis Highway, Suite 400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22202  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/207,169A  
FILING DATE: 08-MAR-1994  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Obion, No. 5674712man F.  
REGISTRATION NUMBER: 24,618  
REFERENCE/DOCKET NUMBER: 2264-061-0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (703) 413-3000  
TELEFAX: (703) 413-2220  
TELEX: 248855 OPAT UR  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 18 amino acids  
TYPE: amino acid  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
FRAGMENT TYPE: internal  
US-08-207-169A-7

Query Match 6.6%; Score 17; DB 1; Length 18;  
Best Local Similarity 94.4%; Pred. No. 1.3;  
Matches 17; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 369 VTSGGSGSGSGSGSGGS 386  
Db 1 VSSGGSGSGSGSGSGGS 18

RESULT 3  
US-08-331-398A-54  
Sequence 54, Application US/08311398A  
Patent No. 5608039  
GENERAL INFORMATION:  
APPLICANT: Pastan, Ira  
APPLICANT: Willingham, Mark  
APPLICANT: Fitzgerald, David  
APPLICANT: Brinkmann, Ulrich  
APPLICANT: Pai, Lee  
TITLE OF INVENTION: Single Chain B3 Antibody Fusion Proteins  
TITLE OF INVENTION: and their uses (as amended)  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Plaza  
City: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331.398A  
FILING DATE: 28-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/767,331  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/596,289  
FILING DATE: 12-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 015280-126110US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-331-398A-54

Query Match 6.2%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SGGSGSGSGSGSGSGGS 386  
Db 1 SGGSGSGSGSGSGSGGS 16

RESULT 4  
US-08-331-397B-54  
Sequence 54, Application US/08311397B  
Patent No. 5981726  
GENERAL INFORMATION:  
APPLICANT: Pastan, Ira  
APPLICANT: Benhar, Itai  
TITLE OF INVENTION: Chimeric and Mutationally Stabilized Tumor-Specific Antibody Fragments, Fusion Proteins, and Uses  
TITLE OF INVENTION: thereof  
NUMBER OF SEQUENCES: 68  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Plaza  
City: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/331.397B  
FILING DATE: 28-OCT-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/767,331  
FILING DATE: 30-SEP-1991  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/596,289  
FILING DATE: 12-OCT-1990  
ATTORNEY/AGENT INFORMATION:  
NAME: Hunter, Tom  
REGISTRATION NUMBER: 38,498  
REFERENCE/DOCKET NUMBER: 015280-126120US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 54:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 16 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-331-397B-54

Query Match 6.2%; Score 16; DB 1; Length 16;  
Best Local Similarity 100.0%; Pred. No. 3.8;  
Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 371 SGGSGSGSGSGSGSGGS 386  
Db 1 SGGSGSGSGSGSGSGGS 16

RESULT 5  
US-08-463-163-5  
Sequence 5, Application US/08463163  
Patent No. 5696237  
GENERAL INFORMATION:  
APPLICANT: Fitzgerald, David J.  
APPLICANT: Chaudhary, Vijay K.  
APPLICANT: Pastan, Ira H.  
APPLICANT: Waldmann, Thomas A.  
APPLICANT: Queen, Cary L.  
TITLE OF INVENTION: Recombinant Antibody-Toxin Fusion Protein  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Tower

CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94105-1492  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/463,163  
 FILING DATE: 05-JUN-1995  
 CLASSIFICATION: 536  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/227,227  
 FILING DATE: 22-JAN-1981  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 06/911,227  
 FILING DATE: 24-SEP-1986  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/341,361  
 FILING DATE: 21-APR-1989  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/865,722  
 FILING DATE: 08-APR-1992  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Weber, Ellen L.  
 REGISTRATION NUMBER: 32,762  
 REFERENCE/DOCKET NUMBER: 015280-12211  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (415) 543-9600  
 TELEFAX: (415) 543-5043  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 17 amino acids  
 TYPE: amino acid  
 STRANDEDNESS:  
 TOPOLOGY: linear  
 MOLECULE TYPE: peptide  
 US-08-463-163-5

Query Match 6.2%; Score 16; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.8;  
 Matches 16; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 371 SGGGSGGGSGGGGS 386  
 Db 1 SGGGSGGGSGGGGS 16

Search completed: September 28, 2005, 10:30:44  
 Job time: 1 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 28, 2005, 10:33:18 ; Search time 0.001 Seconds  
(without alignments)  
304.066 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLTSGAEVKKPKSSSVK.....RYTFGGQTRLEIKLVRGSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapept 0.5

Searched: 17 seqs, 1174 residues

Total number of hits satisfying chosen parameters: 17

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 17 summaries

Database : rapb.subdb.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	88	34.0	95	1 US-10-194-975-84	Sequence 84, Appl
2	88	34.0	95	1 US-10-194-975-85	Sequence 85, Appl
3	88	34.0	95	1 US-10-308-817-31	Sequence 31, Appl
4	88	34.0	95	1 US-10-308-817-32	Sequence 32, Appl
5	88	34.0	95	1 US-10-453-698-31	Sequence 31, Appl
6	88	34.0	95	1 US-10-453-698-32	Sequence 32, Appl
7	88	34.0	95	1 US-10-379-392-95	Sequence 95, Appl
8	88	34.0	95	1 US-10-379-392-96	Sequence 96, Appl
9	88	34.0	95	1 US-10-869-355-22	Sequence 22, Appl
10	88	34.0	95	1 US-10-984-960A-106	Sequence 106, App
11	30	11.6	32	1 US-10-923-068-75	Sequence 75, Appl
12	30	11.6	32	1 US-10-923-068-84	Sequence 84, Appl
13	30	11.6	32	1 US-10-984-960A-34	Sequence 34, Appl
14	30	11.6	32	1 US-10-984-960A-70	Sequence 70, Appl
15	30	11.6	32	1 US-10-901-736-40	Sequence 40, Appl
16	29	11.2	32	1 US-10-901-736-41	Sequence 41, Appl
17	29	11.2	32	1 US-10-901-736-42	Sequence 42, Appl

## ALIGNMENTS

RESULT 1  
US-10-194-975-84  
; Sequence 84, Application US/10194975  
; Publication No. US20030039649A1  
; GENERAL INFORMATION:  
; APPLICANT: Foote, Jefferson  
; TITLE OF INVENTION: Super Humanized Antibodies  
; FILE REFERENCE: 501231.01  
; CURRENT APPLICATION NUMBER: US/10/194, 975  
; CURRENT FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: US 60/305,111

; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 84  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-194-975-84

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELRLTSPATLVSPPERATLSCRASESVSADIAMTQKRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTGSPATLVSPPERATLSCRAQSVSNNIAMTQKRGQAPRLIIYGASTRATGIP 60  
QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 2  
US-10-194-975-85  
; Sequence 85, Application US/10194975  
; Publication No. US20030039649A1  
; GENERAL INFORMATION:  
; APPLICANT: Foote, Jefferson  
; TITLE OF INVENTION: Super Humanized Antibodies  
; FILE REFERENCE: 501231.01  
; CURRENT APPLICATION NUMBER: US/10/194, 975  
; CURRENT FILING DATE: 2002-10-10  
; PRIOR APPLICATION NUMBER: US 60/305,111  
; PRIOR FILING DATE: 2001-07-12  
; NUMBER OF SEQ ID NOS: 122  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 85  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-194-975-85

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELRLTSPATLVSPPERATLSCRASESVSADIAMTQKRGQAPRLIIYGASTRATGVP 446  
DB 1 EIVMTGSPATLVSPPERATLSCRAQSVSNNIAMTQKRGQAPRLIIYGASTRATGIP 60  
QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGTEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 3  
US-10-308-817-31  
; Sequence 31, Application US/10308817  
; Publication No. US20030219861A1  
; GENERAL INFORMATION:  
; APPLICANT: Rothen, Ruessell  
; APPLICANT: Wu, Dayang  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 1087-37  
; CURRENT APPLICATION NUMBER: US/10/308, 817  
; CURRENT FILING DATE: 2002-12-03  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 31  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-308-817-31

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 4  
US-10-308-817-32  
; Sequence 32, Application US/10308817  
; Publication No. US20030219861A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; APPLICANT: Wu, Dayang  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 1087-37  
; CURRENT APPLICATION NUMBER: US/10/308, 817  
; CURRENT FILING DATE: 2002-12-03  
; NUMBER OF SEQ ID NOS: 195  
; SOFTWARE: Patentin version 3.1  
; SEQ ID NO 32  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-308-817-32

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 5  
US-10-453-698-31  
; Sequence 31, Application US/10453698  
; Publication No. US20040038308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 82 CIP (1087-37 CIP)  
; CURRENT APPLICATION NUMBER: US/10/453, 698  
; CURRENT FILING DATE: 2003-06-03  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 31  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-453-698-31

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 6  
US-10-453-698-32  
; Sequence 32, Application US/10453698  
; Publication No. US20040038308A1  
; GENERAL INFORMATION:  
; APPLICANT: Rother, Russell  
; TITLE OF INVENTION: HYBRID ANTIBODIES  
; FILE REFERENCE: 82 CIP (1087-37 CIP)  
; CURRENT APPLICATION NUMBER: US/10/453, 698  
; CURRENT FILING DATE: 2003-06-03  
; NUMBER OF SEQ ID NOS: 196  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 32  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: human  
US-10-453-698-32

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 7  
US-10-379-392-95  
; Sequence 95, Application US/10379392  
; Publication No. US20040110226A1  
; GENERAL INFORMATION:  
; APPLICANT: Lazar, Gregory Alan  
; APPLICANT: Marshall, John Rudolf  
; APPLICANT: Dahlia, Baabli I.  
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION  
; FILE REFERENCE: A-71386-3 463077-236  
; CURRENT APPLICATION NUMBER: US/10/379, 392  
; CURRENT FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360, 843  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 60/384, 197  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: Patentin version 3.2  
; SEQ ID NO 95  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-379-392-95

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;

QY 387 ELETQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGVA 446  
DB 1 EIVMTQSPATLSVSPGERATLSCRASVSDDLAWYQKPGQAPRLLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 481  
DB 61 RFGSGSGAEFTLTISLQSEDFAVYYCCQYNNWP 95

RESULT 8  
US-10-379-392-96  
; Sequence 96, Application US/10379392  
; Publication No. US20040110226A1  
; GENERAL INFORMATION:  
; APPLICANT: Lazar, Gregory Alan  
; APPLICANT: Desjardais, John Rudolf  
; APPLICANT: Marshall, Shannon Alicia  
; APPLICANT: Dahiyat, Basail I.  
; TITLE OF INVENTION: ANTIBODY OPTIMIZATION  
; FILE REFERENCE: A-71386-3 463077-236  
; CURRENT APPLICATION NUMBER: US/10/379,392  
; PRIOR FILING DATE: 2003-03-03  
; PRIOR APPLICATION NUMBER: US 60/360,843  
; PRIOR FILING DATE: 2002-03-01  
; PRIOR APPLICATION NUMBER: US 60/384,197  
; PRIOR FILING DATE: 2002-05-29  
; NUMBER OF SEQ ID NOS: 184  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 96  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-379-392-96

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 387 EELVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 446  
RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
Db 1 RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 9  
US-10-869-355-22  
; Sequence 22, Application US/10869355  
; Publication No. US20050048578A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhang, Dongxiao  
; TITLE OF INVENTION: METHODS OF SCREENING FOR MONOCLONAL  
; FILE REFERENCE: EPT-007  
; CURRENT APPLICATION NUMBER: US/10/869,355  
; PRIOR FILING DATE: 2004-06-15  
; PRIOR APPLICATION NUMBER: 60/483,391  
; PRIOR FILING DATE: 2003-06-26  
; PRIOR APPLICATION NUMBER: 60/484,185  
; PRIOR FILING DATE: 2003-06-30  
; NUMBER OF SEQ ID NOS: 38  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 22  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Oryctolagus cuniculus  
US-10-869-355-22

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 387 EELVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 446  
RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
Db 1 RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 10  
US-10-984-960A-106  
; Sequence 106, Application US/10984960A  
; Publication No. US20050142137A1  
; GENERAL INFORMATION:  
; APPLICANT: Gallo, Michael  
; APPLICANT: Chui, Daniel  
; APPLICANT: Zhong, Haibong  
; APPLICANT: Ara, Gulshan  
; APPLICANT: Laroche, William J.  
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR  
; FILE REFERENCE: Cura 970  
; CURRENT APPLICATION NUMBER: US/10/984,960A  
; PRIOR FILING DATE: 2004-11-08  
; PRIOR APPLICATION NUMBER: 60/518,275  
; PRIOR FILING DATE: 2003-11-07  
; NUMBER OF SEQ ID NOS: 112  
; SOFTWARE: CuraSeqlet version 0.1  
; SEQ ID NO 106  
; LENGTH: 95  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-984-960A-106

Query Match 34.0%; Score 88; DB 1; Length 95;  
Best Local Similarity 92.6%; Pred. No. 3;  
Matches 88; Conservative 0; Mismatches 7; Indels 0; Gaps 0;  
Db 1 EIWVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 60  
Qy 387 EELVTSQPATLSVSPGERATLSCRASQSVSNLAWYQKQKQAPRLIIYGASTRATGIPA 446  
RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 481  
Db 1 RFSGSGSGAEFTLTISLQSEDFAVYYCCQYNNMP 95

RESULT 11  
US-10-923-068-75  
; Sequence 75, Application US/10923068  
; Publication No. US20050042664A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Herren  
; APPLICANT: Dameschroder, Melissa  
; TITLE OF INVENTION: HUMANIZATION OF ANTIBODIES  
; FILE REFERENCE: AB600US  
; CURRENT APPLICATION NUMBER: US/10/923,068  
; PRIOR FILING DATE: 2004-08-20  
; NUMBER OF SEQ ID NOS: 518  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 75  
; LENGTH: 32  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-923-068-75

Query Match 11.6%; Score 30; DB 1; Length 32;  
Best Local Similarity 93.8%; Pred. No. 16;  
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Db 1 GIPARFSGSGAEFTLTISLQSEDFAVYYC 32  
Qy 443 GIPARFSGSGAEFTLTISLQSEDFAVYYC 474  
RFSGSGSGAEFTLTISLQSEDFAVYYC 481  
Db 1 RFSGSGSGAEFTLTISLQSEDFAVYYC 95

RESULT 12  
US-10-923-068-84  
; Sequence 84, Application US/10923068  
; Publication No. US20050042664A1  
; GENERAL INFORMATION:  
; APPLICANT: Wu, Herren  
US-10-923-068-84

```

; APPLICANT: Dall'Acqua, William
; APPLICANT: Damechroder, Melissa
; TITLE OF INVENTION: HUMANIZATION OF ANTIBODIES
; FILE REFERENCE: AB600US
; CURRENT APPLICATION NUMBER: US/10/923,068
; CURRENT FILING DATE: 2004-08-20
; NUMBER OF SEQ ID NOS: 518
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 84
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-923-068-84

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSSEDPFAYYC 32

RESULT 13
US-10-984-960A-34
; Sequence 34, Application US/10984960A
; Publication No. US20050142137A1
; GENERAL INFORMATION:
; APPLICANT: Gallo, Michael
; APPLICANT: Chui, Daniel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ara, Gulshan
; APPLICANT: Larochelle, William J.
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR
; FILE REFERENCE: Cura 970
; CURRENT APPLICATION NUMBER: US/10/984,960A
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: 60/518,275
; PRIOR FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 112
; SOFTWARE: Curaseq1st version 0.1
; SEQ ID NO 34
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-960A-34

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSSEDPFAYYC 32

RESULT 14
US-10-984-960A-70
; Sequence 70, Application US/10984960A
; Publication No. US20050142137A1
; GENERAL INFORMATION:
; APPLICANT: Gallo, Michael
; APPLICANT: Chui, Daniel
; APPLICANT: Zhong, Haihong
; APPLICANT: Ara, Gulshan
; APPLICANT: Larochelle, William J.
; TITLE OF INVENTION: ANTIBODIES AGAINST SECRETORYLEUKOCYTE PROTEASE INHIBITOR
; FILE REFERENCE: Cura 970
; CURRENT APPLICATION NUMBER: US/10/984,960A
; CURRENT FILING DATE: 2004-11-08
; PRIOR APPLICATION NUMBER: 60/518,275
; PRIOR FILING DATE: 2003-11-07
; NUMBER OF SEQ ID NOS: 112
```

```

; SOFTWARE: Curaseq1st version 0.1
; SEQ ID NO 70
; LENGTH: 32
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-984-960A-70

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSSEDPFAYYC 32

RESULT 15
US-10-901-736-40
; Sequence 40, Application US/10901736
; Publication No. US20050169909A1
; GENERAL INFORMATION:
; APPLICANT: TANOX, INC.
; APPLICANT: SINGH, Sanjaya
; APPLICANT: HUANG, Danyang
; APPLICANT: FUNG, Sek Chung
; TITLE OF INVENTION: Identification of Unique, High Affinity Ige Epitopes
; FILE REFERENCE: TNX-1030
; CURRENT APPLICATION NUMBER: US/10/901,736
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 60/444,229
; PRIOR FILING DATE: 2004-02-01
; PRIOR APPLICATION NUMBER: PCT/US04/02892
; PRIOR FILING DATE: 2004-02-02
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 40
; LENGTH: 32
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: FRU3 VARIANT 136
US-10-901-736-40

Query Match      11.6%; Score 30; DB 1; Length 32;
Best Local Similarity 93.8%; Pred. No. 16;
Matches 30; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      443 GVPARFSGSGGAEFTLTITSSLOSSEDPFAYYC 474
DB      1 GIPARFSGSGGTEFTLTITSSLOSSEDPFAYYC 32

RESULT 16
US-10-901-736-41
; Sequence 41, Application US/10901736
; Publication No. US20050169909A1
; GENERAL INFORMATION:
; APPLICANT: TANOX, INC.
; APPLICANT: SINGH, Sanjaya
; APPLICANT: HUANG, Danyang
; APPLICANT: FUNG, Sek Chung
; TITLE OF INVENTION: Identification of Unique, High Affinity Ige Epitopes
; FILE REFERENCE: TNX-1030
; CURRENT APPLICATION NUMBER: US/10/901,736
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 60/444,229
; PRIOR FILING DATE: 2004-02-01
; PRIOR APPLICATION NUMBER: PCT/US04/02892
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: PCT/US04/02894
; PRIOR FILING DATE: 2004-02-02
```

```

      : NUMBER OF SEQ ID NOS:77
      : SOFTWARE: PatentIn version 3.2
      : SEQ ID NO 41
      : LENGTH: 32
      : TYPE: PRt
      : ORGANISM: ARTIFICIAL
      : FEATURE:
      : OTHER INFORMATION: FR13 VARIANT 1
      : US-10-901-736-41

```

Query Match	11.2%;	Score 29;	DB 1;	Length 32
Best Local Similarity	90.6%;	Pred. No. 16;		
Matches 29; Conservative	0;	Mismatches	3;	Indels

```
QY      443 GVPARFSGSGSGAEFTLTISLSQSEDAFVAYYC 474
          | | | | | | | | | | | | | | | | | |
Db      1 GIPSRFSGSGSGTEFTLTISLSQSEDAFVAYYC 32
```

```

RESULT 17
US-10-901-736-42
; Sequence 42, Application US/10901736
; Publication No. US2005016909A1
; GENERAL INFORMATION:
; APPLICANT: TANOX, INC.
; APPLICANT: SINGH, Sanjaya
; APPLICANT: HUANG, Danyang
; APPLICANT: FONG, Sek Chung
; TITLE OF INVENTION: Identification of Unique, High Affinity IGE Epitopes
; FILE REFERENCE: TNX-1030
; CURRENT APPLICATION NUMBER: US/10/901, 736
; CURRENT FILING DATE: 2004-07-29
; PRIOR APPLICATION NUMBER: 60/444, 229
; PRIOR FILING DATE: 2003-02-01
; PRIOR APPLICATION NUMBER: PCT/US04/02892
; PRIOR FILING DATE: 2004-02-02
; PRIOR APPLICATION NUMBER: PCT/US04/02894
; NUMBER OF SEQ ID NOS: 77
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 42
; LENGTH: 32
; TYPE: PRT
; ORGANISM: ARTIFICIAL
; FEATURE:
; OTHER INFORMATION: FRL3 VARIANT 13
US-10-901-736-42

```

Query Match	11.2%	Score 29;	DB 1;	Length 32;
Best Local Similarity	90.6%	Pred. NO. 16;		
Matches 29;	Conservative 0;	Mismatches 3;	Indels 0;	Gaps 0;

```
QY      443 GVPARFSGSGSSEFTLTISLSLQSDPAVYYC 474
          |||||
Db       1 GIPIRFGSGSGSEFTLTISLSLQSDPADYYC 32
```

Search completed: September 28, 2005, 10:33:18  
Job time : 0.001 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 28, 2005, 10:42:58 ; Search time 0.001 Seconds  
(without alignments)  
21.756 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 259  
Sequence: 1 QVQLSGAEVKKKSGSVKV.....RYFGQSTRLEIKLVPRSG 259

Scoring table: UNITARY AA  
Gapop 10.0 , Gapext 0.5

Searched: 1 seqs, 84 residues

Total number of hits satisfying chosen parameters: 1

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 100 summaries

Database : tpr.subdb.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	80	30.9	84	1 S34099	

## ALIGNMENTS

RESULT 1  
S34099

Ig kappa chain V region - human  
C:Species: Homo sapiens (man)  
C>Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 09-Jul-2004  
C:Accession: S34099  
R:Wagner, S.D.; Luzatco, L.  
Eur. J. Immunol. 23, 391-397, 1993  
A:Title: V-kappa gene segments rearranged in chronic lymphocytic leukemia are distributed  
A:Reference number: S34076; MUID:93170387; PMID:8436174  
A:Accession: S34099  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-84 <WAG>  
A:Cross-references: UNIPROT:Q9UL83; EMBL:X67183  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
P:5-79/Domain: immunoglobulin homology <IMM>

Query Match 30.9%; Score 80; DB 1; Length 84;  
Best Local Similarity 95.2%; Pred. No. 0;  
Matches 80; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 398 SVSPGERATISCRASESVSDLAHYQQKPGQAPRLIYGASTRATGVPAFFSGSGGAFF 457  
Db 1 SVSPGERATISCRASESVSDLAHYQQKPGQAPRLIYGASTRATGVPAFFSGSGGAFF 60

QY 458 TLTISSLOSEDFAVYYCQGYNNMP 481  
Db 61 TLTISSLOSEDFAVYYCQGYNNMP 84

Search completed: September 28, 2005, 10:42:59  
Job time : 0.001 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 09:37:05 ; Search time 140 Seconds  
(without alignments)  
7310.971 Million cell updates/sec

Title: US-09-936-702-4

1440

Perfect score: 1 atgacccggggagtcctt.....gtcagcagtcacataactg 1440

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : Issued Patents NA.\*

1: /cgn2\_6/ptodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/ptodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/ptodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/ptodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/ptodata/1/ina/PCTUS.COMB.seq:\*  
6: /cgn2\_6/ptodata/1/ina/backfiles1.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	625	43.4	1213	6	5223418-1
2	624.4	43.4	1415	5	PCT-US92-08090-4
3	624.4	43.4	1421	5	PCT-US92-08090-3
4	624.4	43.4	1448	5	PCT-US92-08090-2
5	624.4	43.4	1742	3	US-09-517-605-7
6	624.4	43.4	1742	4	US-09-023-655-1013
7	624.4	43.4	2465	5	PCT-US92-08090-1
8	622.8	43.2	1304	2	US-08-284-391B-28
9	622.8	43.2	1304	3	US-09-218-950-28
10	622.8	43.2	1304	4	US-08-394-388A-28
11	622.8	43.2	1389	2	US-08-417-495-2
12	622.8	43.2	1389	3	US-09-218-950-2
13	622.8	43.2	1389	2	US-08-284-391B-2
14	622.8	43.2	1389	4	US-08-394-388A-2
15	622.8	43.2	1389	5	PCT-US92-01785-2
16	622.8	43.2	1389	5	PCT-US95-00454-2
17	622.8	43.2	1599	2	US-08-417-495-3
18	622.8	43.2	1599	2	US-08-284-391B-3
19	622.8	43.2	1599	4	US-09-218-950-3
20	622.8	43.2	1599	3	US-08-394-388A-3
21	622.8	43.2	1599	5	PCT-US92-01785-3
22	622.8	43.2	1599	5	PCT-US95-00454-3
23	622.8	43.2	1728	2	US-08-417-495-1
24	622.8	43.2	1728	2	US-08-284-391B-1
25	622.8	43.2	1728	3	US-09-218-950-1
26	622.8	43.2	1728	4	US-08-394-388A-1
27	622.8	43.2	1728	5	PCT-US92-01785-1

28	622.8	43.2	1728	5	PCT-US95-00454-1	Sequence 1, Appli
29	622.8	43.2	1742	3	US-08-466-368-3	Sequence 3, Appli
30	622.8	43.2	1742	4	US-08-470-998-1	Sequence 1, Appli
31	622.8	43.2	1742	4	US-08-328-500-8	Sequence 8, Appli
32	622.8	43.2	2589	4	US-08-472-888A-5	Sequence 5, Appli
33	621.2	43.1	1273	3	US-08-466-368-1	Sequence 1, Appli
34	621.2	43.1	1273	4	US-08-328-500-1	Sequence 1, Appli
35	621.2	43.1	1416	1	US-08-236-311-2	Sequence 2, Appli
36	621.2	43.1	1416	3	US-08-457-918-2	Sequence 2, Appli
37	621.2	43.1	1416	4	US-10-157-408-2	Sequence 2, Appli
38	621.2	43.1	2940	2	US-08-418-848A-3	Sequence 3, Appli
39	621.2	43.1	3426	2	US-08-418-848A-2	Sequence 2, Appli
40	621.2	43.1	3480	2	US-08-418-848A-4	Sequence 4, Appli
41	621.2	43.1	3721	2	US-08-418-848A-5	Sequence 5, Appli
42	621.2	43.1	3993	2	US-08-418-848A-6	Sequence 6, Appli
43	621.2	43.1	4059	2	US-08-418-848A-7	Sequence 7, Appli
44	621.2	43.1	4632	2	US-08-418-848A-8	Sequence 8, Appli
45	619.6	43.0	1009	6	5223394-8	Patent No. 5223394

## ALIGNMENTS

RESULT 1									
5223418-1									
; Patent No. 5223418									
; APPLICANT: ARCURI, EDWARD J.; BRAMNER, MARY E.; DONOVAN, MARY									
; J.; GERBER, ROBERT G.; KELLER, JOHN A.									
; TITLE OF INVENTION: METHOD OF IMPROVING THE YIELD OF									
; HETEROLOGOUS PROTEINS PRODUCED BY STREPTOMYCES LIVIDANS									
; NUMBER OF SEQUENCES: 2									
; CURRENT APPLICATION DATA:									
; APPLICATION NUMBER: US/07/589, 979									
; FILING DATE: 28-SEP-1990									
; SEQ ID NO.1:									
; LENGTH: 1213									
5223418-1									
Query Match 43.4%; Score 625; DB 6; Length 1213;									
Best Local Similarity 97.7%; Pred. No. 3.8e-138;									
Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;									
QY	1	ATGAACCGGGAGTCCCTTTAGGCACTTCTTGTGCTGCACTGGCCCTCTCCCA	60						
DB	79	ATGAACCGGGAGTCCCTTTAGGCACTTCTTGTGCTGCACTGGCCCTCTCCCA	138						
QY	61	GCAGCACTCAGGGAAGAAAGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC	120						
DB	139	GCAGCACTCAGGGAAGAAAGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC	198						
QY	121	TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATAAATCCCAACAGATAAG	180						
DB	199	TGTACAGCTTCCAGAAAGAGCATCAATTCACCTGGAATAAATCCCAACAGATAAG	258						
QY	181	ATTCTGGGAATCAGGGCTCTTCTTAATCAAGTGCATCCAGCTGAATGCGCT	240						
DB	259	ATTCTGGGAATCAGGGCTCTTCTTAATCAAGTGCATCCAGCTGAATGCGCT	318						
QY	241	GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG	300						
DB	319	GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG	378						
QY	301	ATGAAGATCAGATTAATTAATTTGTGAAGTGAGAGCAAGAAAGAGGTGCAATTG	360						
DB	379	ATGAAGATCAGATTAATTAATTTGTGAAGTGAGAGCAAGAAAGAGGTGCAATTG	438						
QY	361	CTAATGTTCCGATTAATTAATTTGTGAAGTGAGAGCAAGAAAGAGGTGCAATTG	420						
DB	439	CTAATGTTCCGATTAATTAATTTGTGAAGTGAGAGCAAGAAAGAGGTGCAATTG	498						
QY	421	CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT	480						
DB	499	CTGACCTTGGAGAGCCCTCTGTAGTACCCCTCAGTGCATGTAGAGTCCAAAGGGGT	558						

QY 481 AAAACATACAGGGGGGAGAACCTCTCCGTGCTCAGCTGGAGCTCCAGATAGTGGC 540  
DB 559 AAAAACATACAGGGGGGAGAACCTCTCCGTGCTCAGCTGGAGCTCCAGATAGTGGC 618  
QY 541 ACCTGACATGACCTGCTTTCAGAACCAAGAGAGTGGAGTTCAAAATAGACATGCTG 600  
DB 619 ACCTGACATGACCTGCTTTCAGAACCAAGAGAGTGGAGTTCAAAATAGACATGCTG 678  
QY 601 GTGCTAGCTTCCAGAGGGCTCCGAGGTGGCGGTGCTGGGGAGGCG 649  
DB 679 GTGCTAGCTTCCAGAGGGCTCCGAGGTGGCGGTGCTGGGGAGGCG 727

## RESULT 2

PCT-US92-08090-4  
Sequence 4, Application PC/TUS9208090  
GENERAL INFORMATION:  
APPLICANT: Tang, J. N.  
TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
TITLE OF INVENTION: Aids  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08090  
FILING DATE: 19920922  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.  
REGISTRATION NUMBER: 31,284  
REFERENCE/DOCKET NUMBER: OMRP129  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 404-572-6508  
TELEFAX: 404-572-6555  
INFORMATION FOR SEQ. ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1415 base pairs  
TYPE: NUCLEIC ACID  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Homo sapien  
TISSUE TYPE: epithelial  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1..6  
OTHER INFORMATION: /note= "Restriction site"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1275..1280  
OTHER INFORMATION: /note= "Restriction site"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 1410..1415  
OTHER INFORMATION: /note= "Restriction site"  
PCT-US92-08090-4

Query Match 43.4%; Score 624.4; DB 5; Length 1415;  
Best Local Similarity 99.8%; Pred. No. 5.5e-138;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGACCGGGAGTCCCTTTTATGCACTTGCTGTGCTGCAAGTGGCGCTCTCCCA 60  
DB 81 ATGACCGGGAGTCCCTTTTATGCACTTGCTGTGCTGCAAGTGGCGCTCTCCCA 140  
QY 61 GCAGCCACTCAGGAAAGAAAGTGTGTGCGCAAAAAGGGATACATGAACTGACC 120  
DB 141 GCAGCCACTCAGGAAAGAAAGTGTGTGCGCAAAAAGGGATACATGAACTGACC 200  
QY 121 TGTACAGCTTCCAGAAAGACATCAATTCATCTGCAAAAATCTCAACCGATTAAG 180  
DB 201 TGTACAGCTTCCAGAAAGACATCAATTCATCTGCAAAAATCTCAACCGATTAAG 260  
QY 181 ATTCTGGAAATCAGGGCTCCTCTTAACTAAAGTCTCAAGCTGAATGATCGGCT 240  
DB 261 ATTCTGGAAATCAGGGCTCCTCTTAACTAAAGTCTCAAGCTGAATGATCGGCT 320  
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 321 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 380  
QY 301 ATGAAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGTGCATTG 360  
DB 381 ATGAAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGTGCATTG 440  
QY 361 CTAGTGTGGATGTAAGTGTGCACTGTGACACCCACTGCTTGAAGGGGCAAGCCTGACC 420  
DB 441 CTAGTGTGGATGTAAGTGTGCACTGTGACACCCACTGCTTGAAGGGGCAAGCCTGACC 500  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATATGAGTCCAGGGGT 480  
DB 501 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATATGAGTCCAGGGGT 560  
QY 481 AAAAACATACAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 561 AAAAACATACAGGGGGGAGAACCTCTCTCGTGTCTCAGCTGAGCTCCAGATAGTGGC 620  
QY 541 ACCTGACATGACCTGCTTTCAGAACCAAGAGAGTGGAGTTCAAAATAGACATGCTG 600  
DB 621 ACCTGACATGACCTGCTTTCAGAACCAAGAGAGTGGAGTTCAAAATAGACATGCTG 680  
QY 601 GTGCTAGCTTCCAGAGGGCTCCGG 626  
DB 681 GTGCTAGCTTCCAGAGGGCTCCAG 706

## RESULT 3

PCT-US92-08090-3  
Sequence 3, Application PC/TUS9208090  
GENERAL INFORMATION:  
APPLICANT: Tang, J. N.  
TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
TITLE OF INVENTION: Aids  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kilpatrick & Cody  
STREET: 100 Peachtree Street  
CITY: Atlanta  
STATE: Georgia  
COUNTRY: U.S.  
ZIP: 30303  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: PCT/US92/08090  
FILING DATE: 19920922  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Pabst, Patrea L.

```

;
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1421 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1416..1421
; OTHER INFORMATION: /note= "Restriction site"
;
PCT-US92-08090-3

```

```

Query Match      43.4%; Score 624.4; DB 5; Length 1421;
Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCAATGGGCGCTCCCTCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCAATGGGCGCTCCCTCCA 140
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGAC 120
DB 141 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGAC 200
QY 121 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGAAAAAAGCTCCCAAGATTAAG 180
DB 201 TGTACAGCTTCCCAAGAAAGAGCATATTCATCTGAAAAAAGCTCCCAAGATTAAG 260
QY 181 ATTCTGGGAATCAGGGGCTCTTCTTAACCTAAAGTCCATCCAAAGCTGAATGCGCGT 240
DB 261 ATTCTGGGAATCAGGGGCTCTTCTTAACCTAAAGTCCATCCAAAGCTGAATGCGCGT 320
QY 241 GACTCAAGAGAAGCCTTTTGGGCAAGAAATTCCTCCCTGATCATCAAGATCTTAA 300
DB 321 GACTCAAGAGAAGCCTTTTGGGCAAGAAATTCCTCCCTGATCATCAAGATCTTAA 380
QY 301 ATAGAAGATTCAGATCTTACATCTGTGAAGTGAAGAGCAAGAGAGAGAGTCAATTG 360
DB 381 ATAGAAGATTCAGATCTTACATCTGTGAAGTGAAGAGCAAGAGAGAGAGTCAATTG 440
QY 361 CTAGGTTGGGATTGATCTGCAACTGTGACACCCACCTGCTCAGGGGAGAGCCGAGC 420
DB 441 CTAGGTTGGGATTGATCTGCAACTGTGACACCCACCTGCTCAGGGGAGAGCCGAGC 500
QY 421 CTGACCTTGAAGAGCCCTGCTGATAGTACCCCTCAGTGCATATGATGAGTCCAGGGT 480
DB 501 CTGACCTTGAAGAGCCCTGCTGATAGTACCCCTCAGTGCATATGATGAGTCCAGGGT 560
QY 481 AAAAAATATCAGGGGGGAAAGACCTCTCGGTCTCAGCTGAGAGCTCCAGATAGTGC 540
DB 561 AAAAAATATCAGGGGGGAAAGACCTCTCGGTCTCAGCTGAGAGCTCCAGATAGTGC 620
QY 541 AACTGAGATGACCTGCTTGGAGAGCAAGAGAAAGTGAAGTCAAAATAGACATCGTG 600

```

```

DB 621 AACTGAGATGACCTGCTTGGAGAGCAAGAGAAAGTGAAGTCAAAATAGACATCGTG 680
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCCG 626
DB 681 GTGCTAGCTTTCCAGAAAGGCTCCCG 706

```

```

RESULT 4
PCT-US92-08090-2
; Sequence 2, Application PC/TUS9208090
; GENERAL INFORMATION:
; APPLICANT: Tang, J. N.
; TITLE OF INVENTION: Fusion Protein Genes for Treatment of
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Kilpatrick & Cody
; STREET: 100 Peachtree Street
; CITY: Atlanta
; STATE: Georgia
; COUNTRY: U.S.
; ZIP: 30303

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US92/08090
; FILING DATE: 19920922
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Pabst, Patrea L.
; REGISTRATION NUMBER: 31,284
; REFERENCE/DOCKET NUMBER: OMRP129
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 404-572-6508
; TELEFAX: 404-572-6555
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1448 base pairs
; TYPE: NUCLEIC ACID
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Homo sapien
; TISSUE TYPE: Epithelial
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1..6
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1275..1280
; OTHER INFORMATION: /note= "Restriction site"
;
; NAME/KEY: misc_feature
; LOCATION: 1444..1448
; OTHER INFORMATION: /note= "Restriction site"
;
PCT-US92-08090-2

```

```

Query Match      43.4%; Score 624.4; DB 5; Length 1448;
Best Local Similarity 99.8%; Pred. No. 5.5e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCAATGGGCGCTCCCTCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGCAATGGGCGCTCCCTCCA 140
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAATGAC 120

```

```
Db 141 GCAGCCACTCAGGAGAAAGTGTGTGGGCAAAAAGGGATACGTGAACTGACC 200
Qy 121 TGTACAGCTTCCCGAAGAAAGCATATTCACCTGAGAAAACCTCCAACTGATAAG 180
Db 201 TGTACAGCTTCCCGAAGAAAGCATATTCACCTGAGAAAACCTCCAACTGATAAG 260
Qy 181 ATTCGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAGTGAATGATCGGCT 240
Db 261 ATTCGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAGTGAATGATCGGCT 320
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
Db 321 GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 380
Qy 301 ATAGAACTCAGATCTTACATCTGTAAGTGAAGGACGAGAAAGGAGGAGTCCAAATTG 360
Db 381 ATAGAACTCAGATCTTACATCTGTAAGTGAAGGACGAGAAAGGAGGAGTCCAAATTG 440
Qy 361 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCAAGGCTGACC 420
Db 441 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCAAGGCTGACC 500
Qy 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGTAGAGAGTCCAAAGGGGT 480
Db 501 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGTAGAGAGTCCAAAGGGGT 560
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGGCTCCAGATAGTGGC 540
Db 561 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGGCTCCAGATAGTGGC 620
Qy 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAATAGACATGCTG 600
Db 621 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAATAGACATGCTG 680
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db 681 GTGCTAGCTTTCAGAAAGGCTCCAG 706
```

```
RESULT 5
US-09-517-605-7
; Sequence 7, Application US/09517605
; Patent No. 6391567
; GENERAL INFORMATION:
; APPLICANT: Liltman, Dan R.
; APPLICANT: Kwon, Douglas S.
; APPLICANT: van Kooyk, Yvette
; APPLICANT: Geljtenbeck, Theo
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO
; FILE REFERENCE: 1049-1-017
; CURRENT APPLICATION NUMBER: US/09/517,605
; NUMBER OF SEQ ID NOS: 17
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1742
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-517-605-7
```

```
Query Match 43.4%; Score 624.4; DB 3; Length 1742;
Best Local Similarity 99.8%; Pred. No. 5.8e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGAACCGGGAGTCCCTTTAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 60
Db 76 ATGAACCGGGAGTCCCTTTAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 135
Qy 61 GCAGCCACTCAGGAGAAAGTGTGTGGGCAAAAAGGGATACGTGAACTGACC 120
Db 136 GCAGCCACTCAGGAGAAAGTGTGTGGGCAAAAAGGGATACGTGAACTGACC 195
```

```
Qy 121 TGTACAGCTTCCCGAAGAAAGCATATTCACCTGAGAAAACCTCCAACTGATAAG 180
Db 196 TGTACAGCTTCCCGAAGAAAGCATATTCACCTGAGAAAACCTCCAACTGATAAG 255
Qy 181 ATTCGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAGTGAATGATCGGCT 240
Db 256 ATTCGGGAAATCAGGGCTCCTTCTTAACCTAAAGTCCATCCAGTGAATGATCGGCT 315
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
Db 316 GACTCAAGAAAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 375
Qy 301 ATAGAACTCAGATCTTACATCTGTAAGTGAAGGACGAGAAAGGAGGAGTCCAAATTG 360
Db 376 ATAGAACTCAGATCTTACATCTGTAAGTGAAGGACGAGAAAGGAGGAGTCCAAATTG 435
Qy 361 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCAAGGCTGACC 420
Db 436 CTAGTGTTCGGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCAAGGCTGACC 495
Qy 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGTAGAGAGTCCAAAGGGGT 480
Db 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGTAGAGAGTCCAAAGGGGT 555
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGGCTCCAGATAGTGGC 540
Db 556 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTCAGCTGAGGCTCCAGATAGTGGC 615
Qy 541 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAATAGACATGCTG 600
Db 616 ACCTGACATGACCTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAATAGACATGCTG 675
Qy 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db 676 GTGCTAGCTTTCAGAAAGGCTCCAG 701
```

```
RESULT 6
US-09-023-655-1013
; Sequence 1013, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Sellhammer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: word Perfect 6.1 for Windows/MS-DOS
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HERewith
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
```

TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1013:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1742 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: 9179143  
 US-09-023-655-1013

Query Match 43.4%; Score 624.4; DB 4; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 5.8e-138;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGACATGCGCGCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGACATGCGCGCTCCCA 135
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACGACC 120
DB 136 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACGACC 195
QY 121 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180
DB 196 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 255
QY 181 ATTCTGGGAATCAGGGCTCCTTTTAATAAGGTCATCCAGGTGAATGCGCGCT 240
DB 256 ATTCTGGGAATCAGGGCTCCTTTTAATAAGGTCATCCAGGTGAATGCGCGCT 315
QY 241 GACTCAAGAGAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAGAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATCTTACATCTGTGAGAGGACCAAGAGAGAGGCTGCAATTG 360
DB 376 ATGAAGACTCAGATCTTACATCTGTGAGAGGACCAAGAGAGAGGCTGCAATTG 435
QY 361 CTAGGTTCCGATTAAGTCCCACTGACACCCCTGCTCAGGGGAGAGCCTGAC 420
DB 436 CTAGGTTCCGATTAAGTCCCACTGACACCCCTGCTCAGGGGAGAGCCTGAC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGTAGAGGAGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGTAGAGGAGGT 555
QY 481 AAAAATATACAGGGGGGAGAAAGCCTCTCTGCTGCTCAGCTGAGCTCCAGGATAGTGC 540
DB 556 AAAAATATACAGGGGGGAGAAAGCCTCTCTGCTGCTCAGCTGAGCTCCAGGATAGTGC 615
QY 541 ACCGTGATGATGACATCTGTGAGAGAACCAAGAGAGGATTTCAAAATACATCTGG 600
DB 616 ACCGTGATGATGACATCTGTGAGAGAACCAAGAGAGGATTTCAAAATACATCTGG 675
QY 601 GTGCTAGCTTTCAGAGAGCCTCCGG 626
DB 676 GTGCTAGCTTTCAGAGAGCCTCCGAG 701

```

RESULT 7  
 PCT-US92-08090-1  
 Sequence 1, Application PC/TUS9208090  
 GENERAL INFORMATION:

APPLICANT: Tang, J. N.  
 TITLE OF INVENTION: Fusion Protein Genes for Treatment of  
 TITLE OF INVENTION: AIDS  
 NUMBER OF SEQUENCES: 15  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Kilpatrick & Cody

```

STREET: 100 Peachtree Street
CITY: Atlanta
STATE: Georgia
COUNTRY: U.S.
ZIP: 30303
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: PCT/US92/08090
FILING DATE: 19920922
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Pabst, Patrea L.
REGISTRATION NUMBER: 31,284
REFERENCE/DOCKET NUMBER: OMRFL29
TELECOMMUNICATION INFORMATION:
TELEPHONE: 404-572-6508
TELEFAX: 404-572-6555
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2465 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Homo sapien
TISSUE TYPE: Epithelial
FEATURE:
NAME/KEY: misc feature
LOCATION: 1..6
OTHER INFORMATION: /note="Restriction site"
NAME/KEY: misc_feature
LOCATION: 2460..2465
OTHER INFORMATION: /note="Restriction site"
PCT-US92-08090-1
Query Match 43.4%; Score 624.4; DB 5; Length 2465;
Best Local Similarity 99.8%; Pred. No. 6.4e-138;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGACATGCGCGCTCCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGACATGCGCGCTCCCA 140
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACGACC 120
DB 141 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACGACC 200
QY 121 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180
DB 201 TGTACAGCTTCCAGAAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 260
QY 181 ATTCTGGGAATCAGGGCTCCTTTTAATAAGGTCATCCAGGTGAATGCGCGCT 240
DB 261 ATTCTGGGAATCAGGGCTCCTTTTAATAAGGTCATCCAGGTGAATGCGCGCT 320
QY 241 GACTCAAGAGAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 321 GACTCAAGAGAAGCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 380
QY 301 ATGAAGACTCAGATCTTACATCTGTGAGAGAACCAAGAGAGAGGCTGCAATTG 360
DB 381 ATGAAGACTCAGATCTTACATCTGTGAGAGAACCAAGAGAGAGGCTGCAATTG 440
QY 361 CTAGGTTCCGATTAAGTCCCACTGACACCCCTGCTCAGGGGAGAGCCTGAC 420

```

Db 441 CTAGTGTTCGATTGACTGCAACTCTGACACCCACTGCTTACGGGGAGAGCCTGACC 500  
Qy 421 CTACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATATGTAGAGTCCAAAGGAGT 480  
Db 501 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATATGTAGAGTCCAAAGGAGT 560  
Qy 481 AAAAACAATACAGGGGGGAGAGACCCCTGCTGTCTCAGCTGAGCTCCAGATATGTC 540  
Db 561 AAAAACAATACAGGGGGGAGAGACCCCTGCTGTCTCAGCTGAGCTCCAGATATGTC 620  
Qy 541 ACCTGACATGCACTGCTTTCGACAAACAGAGAGGTGAGTTCAAATATGACATCTGTG 600  
Db 621 ACCTGACATGCACTGCTTTCGACAAACAGAGAGGTGAGTTCAAATATGACATCTGTG 680  
Qy 601 GTGCTAGCTTTCAGAGAGCCTCCGG 626  
Db 681 GTGCTAGCTTTCAGAGAGCCTCCAG 706

## RESULT 8

US-08-284-391B-28  
; Sequence 28, Application US/08284391B  
; Patent No. 5851828  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Banapour, Babak  
; APPLICANT: Romeo, Charles  
; APPLICANT: Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/284,391B  
; FILING DATE: 02-AUG-1994  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/195,395  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elbing, Karen L.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/247001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 28:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1304 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; US-08-284-391B-28

Query Match 43.2%; Score 622.8; DB 2; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1,36-137;

Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGAGTCCCTTTTATGACATCTCTGAGTGTGCAACTGGGCTCTCCCA 60  
Db 111 ATGAACCGGGAGAGTCCCTTTTATGACATCTCTGAGTGTGCAACTGGGCTCTCCCA 170  
Qy 61 GCAGCCACTCAGGGAAACAAAGTGTGCTGGGCAAAAAAGGGATACATGTAACCTGACC 120  
Db 171 GCAGCCACTCAGGGAAACAAAGTGTGCTGGGCAAAAAAGGGATACATGTAACCTGACC 230  
Qy 121 TGTACGCTTCCAGAAAGACATACATTCATCGGAAAAATCCACCATGATTAAG 180  
Db 231 TGTACGCTTCCAGAAAGACATACATTCATCGGAAAAATCCACCATGATTAAG 290  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 240  
Db 291 ATTCTGGGAATCAGGGCTCTCTTAACTAAAGTCCATCAAGCTGAATGATCGGCT 350  
Qy 241 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 300  
Db 351 GACTCAAGAAAGACCTTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAAG 410  
Qy 301 ATGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 360  
Db 411 ATGAAGACTCAGATACTTACATCTGTGAAGTGAAGACCAAGAGAGGTGCAATTG 470  
Qy 361 CTAGTGTTCGATTGACTGCACTGCACTGACACCACTGCTTACAGGGGACAGGCTGACC 420  
Db 471 CTAGTGTTCGATTGACTGCACTGCACTGACACCACTGCTTACAGGGGACAGGCTGACC 530  
Qy 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 590  
Qy 481 AAAAACAATACAGGGGGGAGAGACCCCTGCTGTCTCAGCTGAGCTCCAGATATGTC 540  
Db 591 AAAAACAATACAGGGGGGAGAGACCCCTGCTGTCTCAGCTGAGCTCCAGATATGTC 650  
Qy 541 ACCTGACATGCACTGCTTTCGAAACCAAGAGAGGTGAGTTCAAATATGACATCTGTG 600  
Db 651 ACCTGACATGCACTGCTTTCGAAACCAAGAGAGGTGAGTTCAAATATGACATCTGTG 710  
Qy 601 GTGCTAGCTTTCAGAGAGCCTCCGG 626  
Db 711 GTGCTAGCTTTCAGAGAGCCTCCAG 736

## RESULT 9

US-09-218-950-28  
; Sequence 28, Application US/09218950  
; Patent No. 6284240  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; APPLICANT: Banapour, Babak  
; APPLICANT: Romeo, Charles  
; APPLICANT: Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: Clark & Elbing LLP  
; STREET: 176 Federal Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: IBM Compatible  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/218,950

FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bibing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-09-218-950-28

Query Match 43.2%; Score 622.8; DB 3; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
DB 171 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCGAGAGAGAGCATATTCACCTGGAATACTCCACAGATAAAG 180  
DB 231 TGTACAGCTTCCGAGAGAGAGCATATTCACCTGGAATACTCCACAGATAAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCCT 350  
QY 241 GACTCAAGAGAGCTTTGGGACCAAGAACTCCCTGATCATCAAGAATCTTAAG 300  
DB 351 GACTCAAGAGAGCTTTGGGACCAAGAACTCCCTGATCATCAAGAATCTTAAG 410  
QY 301 ATAGAAGACTAGATATTAATCTGTGAAGTGAAGACCAAGAGAGAGGTCGAATTG 360  
DB 411 ATAGAAGACTAGATATTAATCTGTGAAGTGAAGACCAAGAGAGAGGTCGAATTG 470  
QY 361 CTAGTGTTCGATTAAGTCACTGCAACCTGCTTCAAGGGCGAAGCTGAC 420  
DB 471 CTAGTGTTCGATTAAGTCACTGCAACCTGCTTCAAGGGCGAAGCTGAC 530  
QY 421 CTGACCTTGAAGAGCCCTGTGTAGTGTGAGTGTGCAATGTGAAGTCCAAAGGGGT 480  
DB 531 CTGACCTTGAAGAGCCCTGTGTAGTGTGAGTGTGCAATGTGAAGTCCAAAGGGGT 590  
QY 481 AAAAATCATAGAGGGGGGAGAGCCCTCTCGGTCTCAGCTGAGCTCAGGATAGTGGC 540  
DB 591 AAAAATCATAGAGGGGGGAGAGCCCTCTCGGTCTCAGCTGAGCTCAGGATAGTGGC 650  
QY 541 ACCTGAGATGACATCTGTCTTGAAGAACCAAGAGAGGTCGAGTTCAAAATAGACATCTGTG 600  
DB 651 ACCTGAGATGACATCTGTCTTGAAGAACCAAGAGAGGTCGAGTTCAAAATAGACATCTGTG 710  
QY 601 GTGCTAGCTTCCAGAGAGGCTCCGG 626

DB 711 GTGCTAGCTTCCAGAGAGGCTCCGG 736

RESULT 10  
US-08-394-388A-28  
Sequence 28, Application US/08394388A  
Patent No. 6753162  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Bibing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Bibing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1304 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-28

Query Match 43.2%; Score 622.8; DB 4; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTAGGCACTTCTGCTGCTGCACTGGCGCTCTCCCA 170  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
DB 171 GCAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAGGGGATACAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCGAGAGAGAGCATATTCACCTGGAATACTCCACAGATAAAG 180  
DB 231 TGTACAGCTTCCGAGAGAGAGCATATTCACCTGGAATACTCCACAGATAAAG 290

```

QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCAAGCTGAATGATCGGCT 240
Db 291 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCAAGCTGAATGATCGGCT 350
QY 241 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 351 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 410
QY 301 ATGAAGACTCAGTACTTATCATCTGTGAAGTGGAGCAAGAGGAGGATGCAATTG 360
Db 411 ATGAAGACTCAGTACTTATCATCTGTGAAGTGGAGCAAGAGGAGGATGCAATTG 470
QY 361 CTAGTGTTCGATTTGACTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
Db 471 CTAGTGTTCGATTTGACTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 530
QY 421 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGATGCAATGAGATCCAAAGGGT 480
Db 531 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGATGCAATGAGATCCAAAGGGT 590
QY 481 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
Db 591 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGGC 650
QY 541 ACCTGACATGCACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATAGACATCTGTG 600
Db 651 ACCTGACATGCACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATAGACATCTGTG 710
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db 711 GTGCTAGCTTTCCAGAAAGGCTCCAG 736

```

## RESULT 11

```

US-08-417-495-2
; Sequence 2, Application US/08417495
; Patent No. 5843728

```

## GENERAL INFORMATION:

```

; APPLICANT: Seed, Brian et al.
; TITLE OF INVENTION: Redirection of Cellular Immunity by
; TITLE OF INVENTION: Receptor
; NUMBER OF SEQUENCES: 27

```

Chimeras

```

; ADDRESSEE: Fish & Richardson
; STREET: 225 Franklin Street
; CITY: Boston

```

```

; STATE: MA

```

```

; COUNTRY: USA

```

```

; ZIP: 02110-2804

```

```

; COMPUTER READABLE FORM:

```

```

; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb

```

```

; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)

```

```

; SOFTWARE: Wordperfect (Version 5.0)

```

```

; CURRENT APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/417,495

```

```

; FILING DATE:

```

```

; CLASSIFICATION: 435

```

```

; PRIOR APPLICATION DATA:

```

```

; APPLICATION NUMBER: US/08/203,866

```

```

; FILING DATE:

```

```

; APPLICATION NUMBER: US/07/847,566

```

```

; FILING DATE:

```

```

; APPLICATION NUMBER: 07/665,961

```

```

; FILING DATE: March 7, 1991

```

```

; CLASSIFICATION: 435

```

```

; ATTORNEY/AGENT INFORMATION:

```

```

; NAME: Clark, Paul T.

```

```

; REGISTRATION NUMBER: 30,162

```

```

; REFERENCE/DOCKET NUMBER: 00786/119002

```

```

; TELECOMMUNICATION INFORMATION:

```

```

; TELEPHONE: (617) 542-5070

```

```

; TELEFAX: (617) 542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1389 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
US-08-417-495-2

```

```

Query Match 43.2%; Score 622.8; DB 2; Length 1389;
Best Local Similarity 99.7%; Pred. No. 1,3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

```

QY 1 ATGACCGGGAGATCCCTTTTAAAGCACTTGCTCTGAGTGTGCAACGAGGCTCTCCCA 60
Db 1 ATGACCGGGAGATCCCTTTTAAAGCACTTGCTCTGAGTGTGCAACGAGGCTCTCCCA 60
QY 61 GCAGCCACTCAGGAGAAAGAAAGTGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
Db 61 GCAGCCACTCAGGAGAAAGAAAGTGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120
QY 121 TGTACAGCTTCCCAAGAAAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 180
Db 121 TGTACAGCTTCCCAAGAAAGACATACATTCACCTGAAAAAATCTCAACGATTAAG 180
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCAAGCTGAATGATCGGCT 240
Db 181 ATTCTGGGAATCAGGGCTCTTCTTAAAGTCAATCAAGCTGAATGATCGGCT 240
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
Db 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGCAAGAGGAGGTGCAATTG 360
Db 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGGAGCAAGAGGAGGTGCAATTG 360
QY 361 CTAGTGTTCGATTTGACTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
Db 361 CTAGTGTTCGATTTGACTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGATGCAATGAGATCCAAAGGGT 480
Db 421 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCACTGATGCAATGAGATCCAAAGGGT 480
QY 481 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
Db 481 AAAACATACAGGGGGGAGAGACCTCTCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
QY 541 ACCTGACATGCACTGTCTTGACAGAACCAAGAAAGGTGAGTTCAAAATAGACATCTGTG 600
Db 541 ACCTGACATGCACTGTCTTGACAGAACCAAGAAAGGTGAGTTCAAAATAGACATCTGTG 600
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626

```

## RESULT 12

```

US-08-284-391B-2
; Sequence 2, Application US/08284391B
; Patent No. 5851828

```

## GENERAL INFORMATION:

```

; APPLICANT: Seed, Brian

```

```

; APPLICANT: Banapour, Babak

```

```

; APPLICANT: Romeo, Charles

```

```

; APPLICANT: Kolanus, Waldemar

```

```

; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED

```

```

; NUMBER OF SEQUENCES: 53

```

```

; CORRESPONDENCE ADDRESS:

```

```

ADDRESSER: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391B
FILING DATE: 02-AUG-1994
CLASSIFICATION: 514
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1389 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-08-284-391B-2

Query Match      43.2%; Score 622.8; DB 2; Length 1389;
Best Local Similarity 99.7%; Pred. No. 1.3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGACTTGTCTTGCTGCACTGGCGCTCTCTCCCA 60
DB 1 ATGAACCGGGAGTCCCTTTAGGACTTGTCTTGCTGCACTGGCGCTCTCTCCCA 60
QY 61 GCAGCCACTCAGAGGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120
DB 61 GCAGCCACTCAGAGGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120
QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCCTGGAATAAACTCCAACTGATTAAG 180
DB 121 TGTACAGCTTCCAGAGAGAGCATACATTCCTGGAATAAACTCCAACTGATTAAG 180
QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCCTGGAATAAACTCCAACTGATTAAG 180
DB 121 TGTACAGCTTCCAGAGAGAGCATACATTCCTGGAATAAACTCCAACTGATTAAG 180
QY 181 ATTCGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 240
DB 181 ATTCGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 240
QY 181 ATTCGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 240
DB 181 ATTCGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATGCGCT 240
QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCCCGCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCCCGCTGATCATCAAGATCTTAAG 300
QY 301 ATAGAGAGCTAGATCTTATCTGTAGAGTGAAGACCAAGAGAGAGGATGCAATTG 360
DB 301 ATAGAGAGCTAGATCTTATCTGTAGAGTGAAGACCAAGAGAGAGGATGCAATTG 360
QY 361 CTAGGTTGGATTTGACTGCACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
DB 361 CTAGGTTGGATTTGACTGCACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
QY 421 CTGACCTTGAAGAGGCGCCCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAGGGGT 480
DB 421 CTGACCTTGAAGAGGCGCCCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAGGGGT 480

```

```

DB 421 CTGACCTTGAAGAGGCGCCCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAGGGGT 480
QY 481 AAAACATACAGGGGGGAGAACCTCTCCGTGTCTGCTGAGCTCCAGATATGTGC 540
DB 481 AAAACATACAGGGGGGAGAACCTCTCCGTGTCTGCTGAGCTCCAGATATGTGC 540
QY 541 ACCTGATCATGCACTGTCTTGAGAGAACAGAGGAGTTCATAAATAGACATCGTG 600
DB 541 ACCTGATCATGCACTGTCTTGAGAGAACAGAGGAGTTCATAAATAGACATCGTG 600
QY 601 GTGCTAGCTTTCAGAGAGGCTCCGG 626
DB 601 GTGCTAGCTTTCAGAGAGGCTCCAG 626

RESULT 13
US-09-218-950-2
Sequence 2, Application US/09218950
Patent No. 6284240
GENERAL INFORMATION:
APPLICANT: Seed, Brian
APPLICANT: Banapour, Babak
APPLICANT: Romeo, Charles
APPLICANT: Kolanus, Waldemar
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED
CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS
NUMBER OF SEQUENCES: 53
CORRESPONDENCE ADDRESS:
ADDRESSEE: Clark & Elbing LLP
STREET: 176 Federal Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02110
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/218,950
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/284,391
FILING DATE: 02-AUG-1994
APPLICATION NUMBER: 08/195,395
FILING DATE: 14-FEB-1994
APPLICATION NUMBER: 07/847,566
FILING DATE: 06-MAR-1992
APPLICATION NUMBER: 07/665,961
FILING DATE: 07-MAR-1991
ATTORNEY/AGENT INFORMATION:
NAME: Elbing, Karen L.
REGISTRATION NUMBER: 35,238
REFERENCE/DOCKET NUMBER: 00786/247001
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-428-0200
TELEFAX: 617-428-7045
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 1389 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: CDNA
US-09-218-950-2

Query Match      43.2%; Score 622.8; DB 3; Length 1389;
Best Local Similarity 99.7%; Pred. No. 1.3e-137;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGTGTGTCGAACTGGCGCTCTCCCA 60  
|  
|  
|  
Db 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGTGTGTCGAACTGGCGCTCTCCCA 60  
|  
|  
|  
QY 61 GCAGCCACTAGAGAAAGAAAGTGTGCTGGCAAAAAGGGATACATGTGAACCTGACC 120  
|  
|  
|  
Db 61 GCAGCCACTAGAGAAAGAAAGTGTGCTGGCAAAAAGGGATACATGTGAACCTGACC 120  
|  
|  
|  
QY 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACATGGAAGAAATCCCAACCAATTAAG 180  
|  
|  
|  
Db 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACATGGAAGAAATCCCAACCAATTAAG 180  
|  
|  
|  
QY 181 ATTCTGGGAATCAAGGCTCTCTTAACTAAAGTCCATCCAGTGAATGATCGGCT 240  
|  
|  
|  
Db 181 ATTCTGGGAATCAAGGCTCTCTTAACTAAAGTCCATCCAGTGAATGATCGGCT 240  
|  
|  
|  
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
|  
|  
|  
Db 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
|  
|  
|  
QY 301 ATGAAGACCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCATTTG 360  
|  
|  
|  
Db 301 ATGAAGACCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCATTTG 360  
|  
|  
|  
QY 361 CTAGTGTTCGGAATGACTGCACTGCAACCCACCTGCTTCAGGGGACAGGCTGACC 420  
|  
|  
|  
Db 361 CTAGTGTTCGGAATGACTGCACTGCAACCCACCTGCTTCAGGGGACAGGCTGACC 420  
|  
|  
|  
QY 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGAGAGTCCAGGGGT 480  
|  
|  
|  
Db 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGAGAGTCCAGGGGT 480  
|  
|  
|  
QY 481 AAAAACAATACAGGGGGGAAAGAACCTCTCGGTGTGAGCTGAGATGATGGC 540  
|  
|  
|  
Db 481 AAAAACAATACAGGGGGGAAAGAACCTCTCGGTGTGAGCTGAGATGATGGC 540  
|  
|  
|  
QY 541 ACCTGACATGACCTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
|  
|  
|  
Db 541 ACCTGACATGACCTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
|  
|  
|  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
|  
|  
|  
Db 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626  
|  
|  
|  
RESULT 14  
US-08-394-388A-2  
Sequence 2, Application US/08394388A  
Patent No. 6753162  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banepour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Ribling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514

QY 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGTGTGTCGAACTGGCGCTCTCCCA 60  
|  
|  
|  
Db 1 ATGAACCGGGAGATCCCTTTTAAAGCACTTCTGTGTGTCGAACTGGCGCTCTCCCA 60  
|  
|  
|  
QY 61 GCAGCCACTAGAGAAAGAAAGTGTGCTGGCAAAAAGGGATACATGTGAACCTGACC 120  
|  
|  
|  
Db 61 GCAGCCACTAGAGAAAGAAAGTGTGCTGGCAAAAAGGGATACATGTGAACCTGACC 120  
|  
|  
|  
QY 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACATGGAAGAAATCCCAACCAATTAAG 180  
|  
|  
|  
Db 121 TGTACAGCTTCCCAAGAAAGAGATCAATTCACATGGAAGAAATCCCAACCAATTAAG 180  
|  
|  
|  
QY 181 ATTCTGGGAATCAAGGCTCTCTTAACTAAAGTCCATCCAGTGAATGATCGGCT 240  
|  
|  
|  
Db 181 ATTCTGGGAATCAAGGCTCTCTTAACTAAAGTCCATCCAGTGAATGATCGGCT 240  
|  
|  
|  
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
|  
|  
|  
Db 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
|  
|  
|  
QY 301 ATGAAGACCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCATTTG 360  
|  
|  
|  
Db 301 ATGAAGACCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTCATTTG 360  
|  
|  
|  
QY 361 CTAGTGTTCGGAATGACTGCACTGCAACCCACCTGCTTCAGGGGACAGGCTGACC 420  
|  
|  
|  
Db 361 CTAGTGTTCGGAATGACTGCACTGCAACCCACCTGCTTCAGGGGACAGGCTGACC 420  
|  
|  
|  
QY 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGAGAGTCCAGGGGT 480  
|  
|  
|  
Db 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAGTGAATGAGAGTCCAGGGGT 480  
|  
|  
|  
QY 481 AAAAACAATACAGGGGGGAAAGAACCTCTCGGTGTGAGCTGAGATGATGGC 540  
|  
|  
|  
Db 481 AAAAACAATACAGGGGGGAAAGAACCTCTCGGTGTGAGCTGAGATGATGGC 540  
|  
|  
|  
QY 541 ACCTGACATGACCTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
|  
|  
|  
Db 541 ACCTGACATGACCTGTCTTGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
|  
|  
|  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
|  
|  
|  
Db 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626  
|  
|  
|  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
US-08-394-388A-2  
Query Match 43.2%; Score 622.8; DB 4; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

RESULT 15  
PCT-US92-01785-2  
; Sequence 2, Application PC/TUS9201785  
; GENERAL INFORMATION:  
; APPLICANT: The General Hospital Corporation  
; TITLE OF INVENTION: Redirection of Cellular Immunity by Receptor  
; TITLE OF INVENTION: Chimeras  
; NUMBER OF SEQUENCES: 27  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson  
; STREET: 225 Franklin Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 502 or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/01785  
; FILING DATE: 19920306  
; CLASSIFICATION: 530  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Clark, Paul T.  
; REGISTRATION NUMBER: 30,162  
; REFERENCE/DOCKET NUMBER: 00786/119002  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ. ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 1389 base pairs  
; TYPE: NUCLEIC ACID  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: DNA (genomic)  
; PCT-US92-01785-2

Query Match 43.2%; Score 622.8; DB 5; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-137;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTGGCACTTGTGCTGTCGCACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGGAGTCCCTTTTGGCACTTGTGCTGTCGCACTGGCGCTCTCCCA 60

QY 61 GCAGCACTCCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
DB 61 GCAGCACTCCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120

QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
DB 121 TGTACAGCTTCCAGAGAGAGCATACATTCACACTGGAAAAAATCCCAACAGATAAAG 180

QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCACACTGGAAAAAATCCCAACAGATAAAG 180  
DB 121 TGTACAGCTTCCAGAGAGAGCATACATTCACACTGGAAAAAATCCCAACAGATAAAG 180

QY 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGAATGCGCGT 240  
DB 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGAATGCGCGT 240

QY 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGAATGCGCGT 240  
DB 181 ATTTGGGAAATCAGGGCTCTTTTAACTAAAGTCCATCCAGGTGAATGCGCGT 240

QY 241 GACTCAGAGAGAGCTTTGGGACCAAGGAATCTCCCTGATCATCAGAACTTTAAG 300  
DB 241 GACTCAGAGAGAGCTTTGGGACCAAGGAATCTCCCTGATCATCAGAACTTTAAG 300

QY 301 ATAGAAGACTCAGATATTACATCTGTGAAGTGAAGACAGAGAGAGAGGTGCAATTG 360  
DB 301 ATAGAAGACTCAGATATTACATCTGTGAAGTGAAGACAGAGAGAGAGGTGCAATTG 360

QY 361 CTAGTGTGGATTGACTGGCCAACTTGACACCCCACTGCTTCAAGGGCAGAGCCTGACC 420  
DB 361 CTAGTGTGGATTGACTGGCCAACTTGACACCCCACTGCTTCAAGGGCAGAGCCTGACC 420

QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 480

QY 481 AAAAATATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAATATACAGGGGGGAGAGACCTCTCCGTCTCAGCTGAGCTCCAGATAGTGGC 540

QY 541 ACCGTGACATGCACTCTCTTGCAGAACCAAGAAAGGTGCACTTCAAAATAGACATCGTG 600  
DB 541 ACCGTGACATGCACTCTCTTGCAGAACCAAGAAAGGTGCACTTCAAAATAGACATCGTG 600

QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626

Search completed: November 22, 2004, 14:15:34  
Job time : 143 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 12:51:30 ; Search time 774 Seconds

(without alignments)  
10053.253 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgaccgggggtccctt.....gtcagcagtcataactcgt 1440

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3627888 seqs, 2701811610 residues

Total number of hits satisfying chosen parameters: 7255776

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:\*

1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq:\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq:\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq:\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq:\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq:\*  
6: /cgn2\_6/ptodata/2/pubpna/PCTUS\_PUBCOMB.seq:\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq:\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq:\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq:\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq:\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq:\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq:\*  
13: /cgn2\_6/ptodata/2/pubpna/US10\_PUBCOMB.seq:\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq:\*  
15: /cgn2\_6/ptodata/2/pubpna/US10C\_PUBCOMB.seq:\*  
16: /cgn2\_6/ptodata/2/pubpna/US10D\_PUBCOMB.seq:\*  
17: /cgn2\_6/ptodata/2/pubpna/US10E\_PUBCOMB.seq:\*  
18: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq:\*  
19: /cgn2\_6/ptodata/2/pubpna/US11\_NEW\_PUB.seq:\*  
20: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq:\*  
21: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	625	43.4	8911	US-10-612-192-3	Sequence 3, Appli
2	624.4	43.4	1377	US-10-103-597A-38	Sequence 38, Appli
3	624.4	43.4	1377	US-10-188-444-38	Sequence 38, Appli
4	624.4	43.4	1742	US-10-151-274-7	Sequence 7, Appli
5	624.4	43.4	1742	US-10-641-643-1013	Sequence 1013, Ap
6	622.8	43.2	1304	US-09-939-537-28	Sequence 28, Appli
7	622.8	43.2	1389	US-09-939-537-2	Sequence 2, Appli
8	622.8	43.2	1389	US-09-243-008-2	Sequence 2, Appli
9	622.8	43.2	1599	US-09-939-537-3	Sequence 3, Appli
10	622.8	43.2	1599	US-09-243-008-3	Sequence 3, Appli
11	622.8	43.2	1728	US-09-939-537-1	Sequence 1, Appli
12	622.8	43.2	1728	US-09-243-008-1	Sequence 1, Appli

13	622.8	43.2	1742	US-09-891-119A-8	Sequence 8, Appli
14	622.8	43.2	1745	US-10-024-329-1	Sequence 1, Appli
15	622.8	43.2	3084	US-10-207-655-169	Sequence 169, App
16	621.2	43.1	1273	US-09-891-119A-1	Sequence 1, Appli
17	621.2	43.1	1416	US-10-157-408-2	Sequence 2, Appli
18	621.2	43.1	1416	US-10-097-044A-2	Sequence 2, Appli
19	621.2	43.1	1416	US-10-769-247-2	Sequence 2, Appli
20	613.4	42.6	1796	US-08-485-163-2	Sequence 2, Appli
21	613.4	42.6	1796	US-09-766-995-1	Sequence 1, Appli
22	612.2	42.5	1149	US-08-485-163-6	Sequence 1, Appli
23	612.2	42.5	1149	US-09-766-995-5	Sequence 5, Appli
24	612.2	42.5	3273	US-10-397-569-7	Sequence 7, Appli
25	612.2	42.5	11228	US-10-397-569-6	Sequence 6, Appli
26	612	42.5	2482	US-08-485-163-4	Sequence 4, Appli
27	612	42.5	2482	US-09-766-995-3	Sequence 3, Appli
28	609	42.3	609	US-10-466-836-24	Sequence 24, Appli
29	607.4	42.2	719	US-09-939-537-30	Sequence 30, Appli
30	571.4	39.7	573	US-10-024-329-31	Sequence 31, Appli
31	552.2	38.3	1508	US-10-157-408-5	Sequence 5, Appli
32	552.2	38.3	1508	US-10-097-044A-5	Sequence 5, Appli
33	552.2	38.3	1508	US-10-769-247-5	Sequence 5, Appli
34	549.4	38.2	1113	US-09-759-841-5	Sequence 5, Appli
35	542.4	37.7	563	US-10-073-118-24	Sequence 24, Appli
36	541.8	37.6	2379	US-10-073-118-25	Sequence 25, Appli
37	535.6	37.2	1769	US-09-934-060A-12	Sequence 12, Appli
38	535.6	37.2	2159	US-09-934-060A-1	Sequence 1, Appli
39	535.6	37.2	2159	US-09-934-060A-5	Sequence 5, Appli
40	534	37.1	534	US-09-934-060A-25	Sequence 25, Appli
41	463.2	32.2	792	US-10-422-628-13	Sequence 13, Appli
42	428.6	29.8	741	US-10-332-673-11	Sequence 9, Appli
43	408.6	28.4	723	US-10-151-882-9	Sequence 10, Appli
44	401	27.8	720	US-10-151-882-10	Sequence 10, Appli
45	401	27.8	900	US-10-203-754A-37	Sequence 37, Appli

## ALIGNMENTS

RESULT 1  
US-10-612-192-3  
; Sequence 3, Application US/10612192  
; Publication No. US20040076636A1  
GENERAL INFORMATION:  
; APPLICANT: Pal, RanaJit  
; APPLICANT: Markham, Phillip  
; APPLICANT: Keen, Timothy  
; APPLICANT: Whitney, Stephen  
; APPLICANT: Kalyanaraman, V.S.  
; TITLE OF INVENTION: HIV Immunogenic Complexes  
; FILE REFERENCE: 00711 CIP  
; CURRENT APPLICATION NUMBER: US/10/612,192  
; PRIOR FILING DATE: 2003-07-02  
; PRIOR APPLICATION NUMBER: US 09/905,962  
; PRIOR FILING DATE: 2001-07-17  
; PRIOR APPLICATION NUMBER: US 09/479,675  
; PRIOR FILING DATE: 2000-01-07  
; PRIOR APPLICATION NUMBER: US 09/075,544  
; NUMBER OF SEQ ID NOS: 3  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 3  
; LENGTH: 8911  
; TYPE: DNA  
; ORGANISM: Artificial  
; FEATURE:  
; OTHER INFORMATION: Expression Vector PTK13+Neo4  
; US-10-612-192-3

Query Match 43.4%; Score 625; DB 16; Length 8911;  
Best Local Similarity 100.0%; Pred. No. 5.6e-171;  
Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 ATGACCGGGAGGCCCTTTAGGACACTTGTCTGTGCTGCACTAGCGCCTCTCCCA 60

```
Db      3815 ATGAACGGGGAGTCCCTTTTAGGCACTTGCTCTGTGTGTCGCAACTGGCGCTCTCCCA 3874
Qy      61  GCAGCCCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC 120
Db      3875 GCAGCCCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC 3934
Qy      121 TGTACAGCTTCCGAGAAAGAGCATCAATTCCTCATGSAAAAACTCCACAGATATAAG 180
Db      3935 TGTACAGCTTCCGAGAAAGAGCATCAATTCCTCATGSAAAAACTCCACAGATATAAG 3994
Qy      181 ATTCTGGGAAATCAGGGCTCTCTTTTAATTAAGGTCCATCCAAAGTGAATGTCGCT 240
Db      3995 ATTCTGGGAAATCAGGGCTCTCTTTTAATTAAGGTCCATCCAAAGTGAATGTCGCT 4054
Qy      241 GACTCAAGAAAGACCTTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 300
Db      4055 GACTCAAGAAAGACCTTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 4114
Qy      301 ATGAAGACTCAGATTAATTAATCTGTGAAGTGGAGGACAGAAAGAGAGTGAATG 360
Db      4115 ATGAAGACTCAGATTAATTAATCTGTGAAGTGGAGGACAGAAAGAGAGTGAATG 4174
Qy      361 CTAGTGTTCGGATTGACTGCACTGTGACACCCACCTGCTTCAGGGGACAGAGCTGACC 420
Db      4175 CTAGTGTTCGGATTGACTGCACTGTGACACCCACCTGCTTCAGGGGACAGAGCTGACC 4234
Qy      421 CTGACCTTGGAGAGCCCCCTGTGTAGTAAAGCCCTCAGTGAAGTGAAGTCAAGGGGT 480
Db      4235 CTGACCTTGGAGAGCCCCCTGTGTAGTAAAGCCCTCAGTGAAGTGAAGTCAAGGGGT 4294
Qy      481 AAAAACAATACAGGGGGGAGAAAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
Db      4295 AAAAACAATACAGGGGGGAGAAAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATAGTGGC 4354
Qy      541 ACCTGACATGACATGTCTTTCAGAACCCAGAAAGGTGAGTTCAAAATATGACATCGTG 600
Db      4355 ACCTGACATGACATGTCTTTCAGAACCCAGAAAGGTGAGTTCAAAATATGACATCGTG 4414
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCG 625
Db      4415 GTGCTAGCTTTCCAGAAAGGCTCCG 4439
```

## RESULT 2

```
US-10-103-597A-38
; Sequence 38, Application US/10103597A
; Publication No. US20030096432A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 992352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1377)
US-10-103-597A-38
```

```
Query Match      43.4%; Score 624.4; DB 14; Length 1377;
Best Local Similarity 99.8%; Pred. No. 4.5e-171;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

```
Qy      1  ATGAACGGGGAGTCCCTTTTAGGCACTTGCTCTGTGTGTCGCAACTGGCGCTCTCCCA 60
Db      1  ATGAACGGGGAGTCCCTTTTAGGCACTTGCTCTGTGTGTCGCAACTGGCGCTCTCCCA 60
Qy      61  GCAGCCCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC 120
Db      61  GCAGCCCTCAGGGAAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGGAACTGACC 120
Qy      121 TGTACAGCTTCCGAGAAAGAGCATCAATTCCTCATGSAAAAACTCCACAGATATAAG 180
Db      121 TGTACAGCTTCCGAGAAAGAGCATCAATTCCTCATGSAAAAACTCCACAGATATAAG 180
Qy      181 ATTCTGGGAAATCAGGGCTCTCTTTTAATTAAGGTCCATCCAAAGTGAATGTCGCT 240
Db      181 ATTCTGGGAAATCAGGGCTCTCTTTTAATTAAGGTCCATCCAAAGTGAATGTCGCT 240
Qy      241 GACTCAAGAAAGACCTTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 300
Db      241 GACTCAAGAAAGACCTTTTGGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAAG 300
Qy      301 ATGAAGACTCAGATTAATTAATCTGTGAAGTGGAGGACAGAAAGAGAGTGAATG 360
Db      301 ATGAAGACTCAGATTAATTAATCTGTGAAGTGGAGGACAGAAAGAGAGTGAATG 360
Qy      361 CTAGTGTTCGGATTGACTGCACTGTGACACCCACCTGCTTCAGGGGACAGAGCTGACC 420
Db      361 CTAGTGTTCGGATTGACTGCACTGTGACACCCACCTGCTTCAGGGGACAGAGCTGACC 420
Qy      421 CTGACCTTGGAGAGCCCCCTGTGTAGTAAAGCCCTCAGTGAAGTGAAGTCAAGGGGT 480
Db      421 CTGACCTTGGAGAGCCCCCTGTGTAGTAAAGCCCTCAGTGAAGTGAAGTCAAGGGGT 480
Qy      481 AAAAACAATACAGGGGGGAGAAAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
Db      481 AAAAACAATACAGGGGGGAGAAAGACCTCTCCGTGTCTGAGCTGAGCTCCAGATAGTGGC 540
Qy      541 ACCTGACATGACATGTCTTTCAGAAACCCAGAAAGGTGAGTTCAAAATATGACATCGTG 600
Db      541 ACCTGACATGACATGTCTTTCAGAAACCCAGAAAGGTGAGTTCAAAATATGACATCGTG 600
Qy      601 GTGCTAGCTTTCCAGAAAGGCTCCG 626
Db      601 GTGCTAGCTTTCCAGAAAGGCTCCG 626
```

## RESULT 3

```
US-10-188-444-38
; Sequence 38, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 992352.1
; PRIOR FILING DATE: 1999-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 1377
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(1377)
US-10-188-444-38
```

```
Query Match      43.4%; Score 624.4; DB 15; Length 1377;
Best Local Similarity 99.8%; Pred. No. 4.5e-171;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
```

QY 1 ATGAACCGGGAGAGTCCCTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGGAGAGTCCCTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACATGGAATCTGACC 120  
DB 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACATGGAATCTGACC 120  
QY 121 TGTACAGCTTCCAGAGAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180  
DB 121 TGTACAGCTTCCAGAGAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180  
QY 181 ATTCGGGAATCAGGGGCTCTTTAACTAAAGGTCATCCAGCTGAATATGCGCT 240  
DB 181 ATTCGGGAATCAGGGGCTCTTTAACTAAAGGTCATCCAGCTGAATATGCGCT 240  
QY 241 GACTCAAGAGAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAGAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
DB 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
QY 361 CTAGTGTTCGAGATGACCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420  
DB 361 CTAGTGTTCGAGATGACCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCCAGGGGT 480  
DB 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCCAGGGGT 480  
QY 481 AAAAAATACAGGGGGGGAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAAATACAGGGGGGGAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCCTGACATGACATGCTCTTTCAGAACCAAGAAAGTGAATTAATACATCTGTG 600  
DB 541 ACCCTGACATGACATGCTCTTTCAGAACCAAGAAAGTGAATTAATACATCTGTG 600  
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626

RESULT 4  
US-10-151-274-7  
Sequence 7, Application US/10151274  
Publication No. US20030064071A1

GENERAL INFORMATION:  
APPLICANT: Littman, Dan R.  
APPLICANT: Kwon, Douglas S.  
APPLICANT: Van Kooyk, Yvette  
APPLICANT: Gelberbeck, Theo  
TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
TITLE OF INVENTION: INTO  
FILE REFERENCE: 1049-1-017  
CURRENT APPLICATION NUMBER: US/10/151,274  
CURRENT FILING DATE: 2002-05-20  
PRIORITY APPLICATION NUMBER: US/09/517,605  
PRIORITY FILING DATE: 2000-03-02  
NUMBER OF SEQ ID NOS: 17  
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 7  
LENGTH: 1742  
TYPE: DNA

ORGANISM: Homo sapiens  
US-10-151-274-7

Query Match 43.4%; Score 624.4; DB 14; Length 1742;  
Best Local Similarity 99.8%; Pred. No. 4,9e-171;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 ATGAACCGGGAGAGTCCCTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 60  
DB 76 ATGAACCGGGAGAGTCCCTTTAGGCACTTGCTGCTGCTGCAACTGGCGCTCTCCCA 135  
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACATGGAATCTGACC 120  
DB 61 GCAGCCACTCAGGGAAGAAAGTGTGTCTGGGCAAAAAAGGGGATACATGGAATCTGACC 120  
QY 121 TGTACAGCTTCCAGAGAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180  
DB 121 TGTACAGCTTCCAGAGAAGAGCATACATTCCTGGAATACTCCACCAATTAAG 180  
QY 181 ATTCGGGAATCAGGGGCTCTTTAACTAAAGTTCATCCAGCTGAATATGCGCT 240  
DB 181 ATTCGGGAATCAGGGGCTCTTTAACTAAAGTTCATCCAGCTGAATATGCGCT 240  
QY 241 GACTCAAGAGAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAGAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
DB 301 ATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
QY 361 CTAGTGTTCGAGATGACCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420  
DB 361 CTAGTGTTCGAGATGACCTGCACTGCAACCCCTGCTTCAAGGGGCAAGGCTGACC 420  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCCAGGGGT 480  
DB 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCCAGGGGT 480  
QY 481 AAAAAATACAGGGGGGGAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAAATACAGGGGGGGAAGACCTCTCTGCTGCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCCTGACATGACATGCTCTTTCAGAACCAAGAAAGTGAATTAATACATCTGTG 600  
DB 541 ACCCTGACATGACATGCTCTTTCAGAACCAAGAAAGTGAATTAATACATCTGTG 600  
QY 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCAGAAAGGCTCCGG 626  
QY 676 GTGCTAGCTTTCAGAAAGGCTCCGG 701  
DB 676 GTGCTAGCTTTCAGAAAGGCTCCGG 701

RESULT 5  
US-10-641-643-1013  
Sequence 1013, Application US/10641643  
Publication No. US20040077003A1

GENERAL INFORMATION:  
APPLICANT: Cooke, Benjamin G.  
APPLICANT: Susan G. Stuart

TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL  
TITLE OF INVENTION: GENE EXPRESSION

NUMBER OF SEQUENCES: 1508  
CORRESPONDENCE ADDRESS:  
ADDRESS: INCYTE PHARMACEUTICALS, INC.  
CITY: PALO ALTO  
STREET: 3174 PORTER DRIVE  
STATE: CALIFORNIA  
COUNTRY: USA

ZIP: 94304  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2

CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/641,643  
FILING DATE: 14-Aug-2003

CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: <Unknown>  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Zeller, Karen J.  
 REGISTRATION NUMBER: 37,071  
 REFERENCE/DOCKET NUMBER: PA-0001 US  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (650) 855-0555  
 TELEFAX: (650) 845-4166  
 INFORMATION FOR SEQ ID NO: 1013:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1742 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 IMMEDIATE SOURCE:  
 LIBRARY: GENBANK  
 CLONE: g179143  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1013 :  
 US-10-641-643-1013

Query Match 43.4%; Score 624.4; DB 16; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 4.9e-171;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGGCGCTCTCCCA 60  
 76 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGGCGCTCTCCCA 135  
 61 GCAGCCACTCAGGGAAGAAAGTGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 120  
 136 GCAGCCACTCAGGGAAGAAAGTGCTGCTGCGCAAAAAGGGGATACAGTGAACCTGACC 195  
 121 TGTACAGCTTCCCGAAGAAAGCATTCATTCACCTGGAAGAACTCCACACGATTAAG 180  
 196 TGTACAGCTTCCCGAAGAAAGCATTCATTCACCTGGAAGAACTCCACACGATTAAG 255  
 181 ATTCTGGGAATCAGGGGCTCTTTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 240  
 256 ATTCTGGGAATCAGGGGCTCTTTTAACTAAAGGTCATCCAGCTGAATGATCGGCT 315  
 241 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
 316 GACTCAAGAAAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
 301 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
 376 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 435  
 361 CTGATGTTGGATTAATGCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 420  
 436 CTGATGTTGGATTAATGCACTGCAACCACTGCTTCAAGGGGCAAGGCTGACC 495  
 421 CTGACCTTGAAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATTAAGAGTCAAGGGGT 480  
 496 CTGACCTTGAAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATTAAGAGTCAAGGGGT 555  
 481 AAAAACAATACAGGGGGGAAAGACCTTCCGTGTCTGAGCTGAGGCTCCAGATAGTGGC 540  
 556 AAAAACAATACAGGGGGGAAAGACCTTCCGTGTCTGAGCTGAGGCTCCAGATAGTGGC 615  
 541 ACCGCGCATGCACTGCTTGGCAAGCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 600  
 616 ACCGCGCATGCACTGCTTGGCAAGCAAGAAAGAGTGAAGTCAAAATAGACATGCTG 675  
 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
 676 GTGCTAGCTTCCAGAGGCTCCAG 701

RESULT 6

US-09-939-537-28  
 Sequence 28, Application US/09939537  
 Publication No. US20030138410A1  
 GENERAL INFORMATION:  
 APPLICANT: Seed, Brian  
 Banepour, Babak  
 Romeo, Charles  
 Kolanus, Waldemar  
 TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
 CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
 NUMBER OF SEQUENCES: 53  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Clark & Elbing LLP  
 STREET: 176 Federal Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/939,537  
 FILING DATE: 24-Aug-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/284,391  
 FILING DATE: 02-AUG-1994  
 APPLICATION NUMBER: 08/195,395  
 FILING DATE: 14-FEB-1994  
 APPLICATION NUMBER: 07/847,566  
 FILING DATE: 06-MAR-1992  
 APPLICATION NUMBER: 07/665,961  
 FILING DATE: 07-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbing, Karen L.  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/247001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-428-0200  
 TELEFAX: 617-428-7045  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 28:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 1304 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 28:

US-09-939-537-28

Query Match 43.2%; Score 622.8; DB 10; Length 1304;  
 Best Local Similarity 99.7%; Pred. No. 1.3e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGGCGCTCTCCCA 60  
 111 ATGAACCGGGGAGTCCCTTTTAGGCACTTGCTGCTGCTGCAACTGGGCGCTCTCCCA 170  
 61 GCAGCCACTCAGGGAAGAAAGTGCTGCGCAAAAAGGGGATACAGTGAAGTGAAC 120  
 171 GCAGCCACTCAGGGAAGAAAGTGCTGCGCAAAAAGGGGATACAGTGAAGTGAAC 230  
 121 TGTACAGCTTCCCGAAGAAAGCATTAATTCACCTGGAAGAACTCCACAGATTAAG 180  
 231 TGTACAGCTTCCCGAAGAAAGCATTAATTCACCTGGAAGAACTCCACAGATTAAG 290  
 181 ATTCTGGGAATCAGGGGCTCTTTTAACTAAAGTTCATCCAGCTGAATGATCGGCT 240  
 291 ATTCTGGGAATCAGGGGCTCTTTTAACTAAAGTTCATCCAGCTGAATGATCGGCT 350

QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAAACTTCCCTCGATTCATCAAGAAATCTTAAG 300  
 Db 351 GACTCAAGAAAGACCTTTGGGACCAAGGAAACTTCCCTCGATTCATCAAGAAATCTTAAG 410  
 QY 301 ATAGAGACTCAGATCTTACATCTGTGAAGTGAGGACAGAGAGGAGAGTGCAATTG 360  
 Db 411 ATAGAGACTCAGATCTTACATCTGTGAAGTGAGGACAGAGAGGAGAGTGCAATTG 470  
 QY 361 CTAGTGTGGGATTGACTGCAACTTGACCACTGCTTCAAGGGGAGAGCTTGACC 420  
 Db 471 CTAGTGTGGGATTGACTGCAACTTGACCACTGCTTCAAGGGGAGAGCTTGACC 530  
 QY 421 CTGACCTTGGAGAGCCCTCGGTACTAGCCCTCAGTGAAGTGAAGTCCAAAGGGT 480  
 Db 531 CTGACCTTGGAGAGCCCTCGGTACTAGCCCTCAGTGAAGTGAAGTCCAAAGGGT 590  
 QY 481 AAAAAATPACAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGGAGCTCCAGATATGTGC 540  
 Db 591 AAAAAATPACAGGGGGGGAAGACCTCTCTCGTGTCTCAGCTGGAGCTCCAGATATGTGC 650  
 QY 541 AACTGGACATGCACTGTCTTGGCAAGACAGAGAGAGGTGAGTTCAAAATGACATCTGTG 600  
 Db 651 AACTGGACATGCACTGTCTTGGCAAGACAGAGAGAGGTGAGTTCAAAATGACATCTGTG 710  
 QY 601 GNGCTAGCTTCCAGAAAGGCTCCGG 626  
 Db 711 GNGCTAGCTTCCAGAAAGGCTTCCAG 736

RESULT 7  
US-09-939-537-2

; Sequence 2, Application US/099339537  
; Publication No. US20030138410A1

```

; GENERAL INFORMATION:
;
; APPLICANT: Seed

```

Banapour, Babak  
Romeo, Charles

1	Kolanus, Waldemar
2	TARGETED CYTOLYSIS OF HIV-INFECTED
3	CELLS BY CHIMERIC CD4 PROTEIN RE

CELLS BI CRITERIAL CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street

CITY: Boston  
STATE: MA

COUNTRY: USA  
ZIP: 02110

COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
;

```

;
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;

```

```

; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 04 MAR 2009

```

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994

APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994

APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/855,061

ATTORNEY/AGENT INFORMATION:  
FILING DATE: 07-MAR-1991  
AFFILIATION NUMBER: 01/0005, 901

NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238

REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:

—

```

? TELEFAX: 617-428-7045
? TELE: <Unknown>
? INFORMATION FOR SEQ ID NO.: 2:
?     SEQUENCE CHARACTERISTICS:
?         LENGTH: 1389 base pairs
?         TYPE: nucleic acid
?         STRANDEDNESS: double
?         TOPOLOGY: linear
?     MOLECULE TYPE: cDNA
?     SEQUENCE DESCRIPTION: SEQ ID NO: 2
US-09-939-537-2

```

Query Match	43.2%;	Score 622.8;	DB 10;	Length 1389;
Best Local Similarity	99.7%;	Pred. No. 1.3e-170;		
Matches 624;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

Oy	1	ATGAACCGGGGAGTCCCTTTTAGGACACTGTCTTGAGTGTGCAAACTGGCCCTCTCCCA	60
Db	1	ATGAACCGGGGAGTCCCTTTTAGGACACTGTCTTGAGTGTGCAAACTGGCCCTCTCCCA	60
Oy	61	GCAGCCACTCAGGGGAAACAAGTGGTGGCTGGGCAAAAAAGGGGATACAGTGAAC	120
Db	61	GCAGCCACTCAGGGGAAACAAGTGGTGGCTGGGCAAAAAAGGGGATACAGTGAAC	120
Oy	121	TGTAACAGCTTCCAGAAAGAGGACATACAAATTCCATGGAATAAATCCAAACCAATAAG	180
Db	121	TGTAACAGCTTCCAGAAAGAGGACATACAAATTCCATGGAATAAATCCAAACCAATAAG	180
Oy	181	ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGGTCATCAAGCTGAATGATGGCT	240
Db	181	ATTCTGGGAAATCAGGGCTCTTCTTAACTAAAGGTCATCAAGCTGAATGATGGCT	240
Oy	241	GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG	300
Db	241	GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG	300
Oy	301	ATGAAGACTCAGATCTTACTCATCTGTGAATGGAGGACGAAGAAGAGAGGTGCAATTG	360
Db	301	ATGAAGACTCAGATCTTACTCATCTGTGAATGGAGGACGAAGAAGAGAGGTGCAATTG	360
Oy	361	CTAGTGTGGATTGACTGCGCAACTGTGACACCCAGCTGCTTCAGGGGCAAGAGCTGACC	420
Db	361	CTAGTGTGGATTGACTGCGCAACTGTGACACCCAGCTGCTTCAGGGGCAAGAGCTGACC	420
Oy	421	CTGAACCTTGAAGAGCCCCCTGTGTGTAGAGCCCCCTCAGTGCAATGTAGAGTCCAAAGGGT	480
Db	421	CTGAACCTTGAAGAGCCCCCTGTGTGTAGAGCCCCCTCAGTGCAATGTAGAGTCCAAAGGGT	480
Oy	481	AAAAACATACAGGGGGGGAAGACCTCTCCGATCTCAGCTGAGAGCTCAAGATATGGG	540
Db	481	AAAAACATACAGGGGGGGAAGACCTCTCCGATCTCAGCTGAGAGCTCCAGATATGGG	540
Oy	541	ACCTGAGACATGACATGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG	600
Db	541	ACCTGAGACATGACATGTCTTTCAGAAACCAAGAAAGGTGAGTTCAAAATAGACATGTG	600
Oy	601	GTGCTAGCTTCCAGAAAGCCCTCCGG	626
Db	601	GTGCTAGCTTCCAGAAAGCCCTCCAG	626

## RESULT 8

US-09-243-008-2  
; Sequence 2, Application US/09243008  
; Publication No. US2004000533A1

GENERAL INFORMATION:

APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
Receptor Chimeras

NUMBER OF SEQUENCES: 40  
CORRESPONDENCE ADDRESS:

ADDRESSEE: Fish & Richardson P.C  
STREET: 225 Franklin Street

CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Mordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/243.008  
FILING DATE: 02-Feb-1999  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/394.176  
FILING DATE: SEPTEMBER 11, 1995  
APPLICATION NUMBER: 08/203.866  
FILING DATE: February 28, 1994  
APPLICATION NUMBER: 07/847.566  
FILING DATE: March 6, 1992  
APPLICATION NUMBER: 07/665.961  
FILING DATE: March 7, 1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Karen F. Lech, Ph.D  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/270001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1389 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: DNA (genomic)  
SEQUENCE DESCRIPTION: SEQ ID NO: 2:  
US-09-243-008-2  
Query Match 43.2%; Score 622.8; DB 11; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 1.3e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCACTGGCGCTCCTCCCA 60  
1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGCTGCACTGGCGCTCCTCCCA 60  
61 GCAGCCACTCAGGGAAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAAC 120  
61 GCAGCCACTCAGGGAAAGAAAGTGTCTGTGGCAAAAAGGGGATACAGTGAAC 120  
121 TGTACACTTCCCGAAGAAAGAGATCAATTCCACTGGAAAAAATCCCAACCAATAAG 180  
121 TGTACACTTCCCGAAGAAAGAGATCAATTCCACTGGAAAAAATCCCAACCAATAAG 180  
181 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAATCAAGTGAATGATCGGCT 240  
181 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAATCAAGTGAATGATGCGGCT 240  
241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
241 GACTCAAGAAAGACCTTTGGGACCAAGAAATTTCCCTGATCATCAAGATCTTAAG 300  
301 ATGAAGACTCAGTACTTACATCTGTGAAGTGAAGGAGCAAGAGAGAGGATCAATTG 360  
301 ATGAAGACTCAGTACTTACATCTGTGAAGTGAAGGAGCAAGAGAGAGGATCAATTG 360  
361 CTAGTGTTCGATGATGCACTGCAACCTGCTTCAAGGGGCAAGAGCTGACC 420  
361 CTAGTGTTCGATGATGCACTGCAACCTGCTTCAAGGGGCAAGAGCTGACC 420  
421 CTGACCTTGAAGAGCCCTCTGTAGTACCCTCAATGATAGAGTCCAAAGGGGT 480  
421 CTGACCTTGAAGAGCCCTCTGTAGTACCCTCAATGATAGAGTCCAAAGGGGT 480

DB 421 CTGACCTTGAAGAGCCCTCTGTAGTACCCTCAATGATAGAGTCCAAAGGGGT 480  
QY 481 AAAACATACAGGGGGGGAAGACCTCTCCGTGTCACTGAGTCCAGATAGTGGC 540  
DB 481 AAAACATACAGGGGGGGAAGACCTCTCCGTGTCACTGAGTCCAGATAGTGGC 540  
QY 541 ACCTGACATGACCTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATGACATGTG 600  
DB 541 ACCTGACATGACCTGTCTTGCAGAACACAGAAAGGTGAGTTCAAAATGACATGTG 600  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTTCCAGAAAGGCTCCAG 626  
RESULT 9  
US-09-939-537-3  
Sequence 3, Application US/09939537  
Publication No. US20030138410A1  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
Banapour, Babak  
Romeo, Charles  
Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
CELLS BY CHIMERIC CD4 RECEPTOR- BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/939,537  
FILING DATE: 24-Aug-2001  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX: <Unknown>  
INFORMATION FOR SEQ ID NO: 3:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1599 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
MOLECULE TYPE: cDNA  
SEQUENCE DESCRIPTION: SEQ ID NO: 3:  
US-09-939-537-3  
Query Match 43.2%; Score 622.8; DB 10; Length 1599;  
Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;



RESULT 11  
US-09-939-537-1  
; Sequence 1, Application US/09939537  
; Publication No. US20030138410A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian  
; Banapur, Babak  
; Romeo, Charles  
; Kolanus, Waldemar  
; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
; NUMBER OF SEQUENCES: 53  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Clark & Ebling LLP  
; STREET: 176 Federal Street  
; City: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/939,537  
; FILING DATE: 24-Aug-2001  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 08/284,391  
; FILING DATE: 02-AUG-1994  
; APPLICATION NUMBER: 08/195,395  
; FILING DATE: 14-FEB-1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: 06-MAR-1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: 07-MAR-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ebling, Karen L.  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/247001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-428-0200  
; TELEFAX: 617-428-7045  
; TELEX: <Unknown>  
; INFORMATION FOR SEQ ID NO: 1:  
; LENGTH: 1728 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-09-939-537-1  
Query Match 43.2%; Score 622.8; DB 10; Length 1728;  
Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 1 ATGAACCGGAGAGTCCCTTTTAGGCACTTGTGCTGCAACTGGCGCTCTCCCA 60  
QY 61 GCAGCCACTAGGGAAGAAAGTGCTGCGCAAAAAGGGGATACAGTGAATGACC 120  
DB 61 GCAGCCACTAGGGAAGAAAGTGCTGCGCAAAAAGGGGATACAGTGAATGACC 120  
QY 121 TGTACAGCTTCCAGAGAAGAGCATATTCCACTGGAATAAATCTCAACAGATAAG 180  
DB 121 TGTACAGCTTCCAGAGAAGAGCATATTCCACTGGAATAAATCTCAACAGATAAG 180  
QY 181 ATTCTGGAAATCAGGGCTCTTCTTAACATAAGATCAGTCAAGTGAATGATGGCT 240  
DB 181 ATTCTGGAAATCAGGGCTCTTCTTAACATAAGATCAGTCAAGTGAATGATGGCT 240

DB 181 ATTCTGGAAATCAGGGCTCTTCTTAACATAAGATCAGTCAAGTGAATGATGGCT 240  
QY 241 GACTCAAGAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 241 GACTCAAGAGAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATGAAGACTCAGATATTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
DB 301 ATGAAGACTCAGATATTACATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
QY 361 CTAGTGTTCGGATTGACTGTGCAACTGTGACACCCACCTGCTTACAGGGGACCTGACC 420  
DB 361 CTAGTGTTCGGATTGACTGTGCAACTGTGACACCCACCTGCTTACAGGGGACCTGACC 420  
QY 421 CTGACCTTGGAGAGCCCTCGTGAAGAGCCCTCAGTGAATGATGAGATCCAAAGGGGT 480  
DB 421 CTGACCTTGGAGAGCCCTCGTGAAGAGCCCTCAGTGAATGATGAGATCCAAAGGGGT 480  
QY 481 AAAAATACAGAGGGGGGAGAACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 540  
DB 481 AAAAATACAGAGGGGGGAGAACCTCTCCGTGCTCAGCTGAGCTCCAGATAGTGGC 540  
QY 541 ACCTGACATGCACTGCTTTCGAGAACCAAGAGAGTGAATTCAAAATGACATGTG 600  
DB 541 ACCTGACATGCACTGCTTTCGAGAACCAAGAGAGTGAATTCAAAATGACATGTG 600  
QY 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626  
DB 601 GTGCTAGCTTCCAGAAAGGCTCCAG 626

RESULT 12  
US-09-243-008-1  
; Sequence 1, Application US/09243008  
; Publication No. US20040005334A1  
; GENERAL INFORMATION:  
; APPLICANT: Seed, Brian et al.  
; TITLE OF INVENTION: Redirection of Cellular Immunity by  
; Receptor Chimeras  
; NUMBER OF SEQUENCES: 40  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Fish & Richardson P.C.  
; STREET: 225 Franklin Street  
; City: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02110-2804  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
; COMPUTER: IBM PS/2 Model 50Z or 55SX  
; OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
; SOFTWARE: Wordperfect (Version 5.0)  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/243,008  
; FILING DATE: 02-Feb-1999  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/394,176  
; FILING DATE: SEPTEMBER 11, 1995  
; APPLICATION NUMBER: 08/203,866  
; FILING DATE: February 28, 1994  
; APPLICATION NUMBER: 07/847,566  
; FILING DATE: March 6, 1992  
; APPLICATION NUMBER: 07/665,961  
; FILING DATE: March 7, 1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Karen F. Lech, Ph.D  
; REGISTRATION NUMBER: 35,238  
; REFERENCE/DOCKET NUMBER: 00786/270001  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (617) 542-5070  
; TELEFAX: (617) 542-8906  
; TELEX: 200154  
; INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:  
 LENGTH: 1728 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: double  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA (genomic)  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-243-008-1

Query Match 43.2%; Score 622.8; DB 11; Length 1728;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTGCTTGTGCTGCTCACTGGCGCTCCCTCCA 60
DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTGCTTGTGCTGCTCACTGGCGCTCCCTCCA 60
QY 61 GCAGCCACTCAGAGGAAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACCTGACC 120
DB 61 GCAGCCACTCAGAGGAAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACCTGACC 120
QY 121 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCACCAATTAAG 180
DB 121 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCACCAATTAAG 180
QY 181 ATTCTGGGAATCAGGGCTCTTTTAATAAGGTCATCCAGCTGAATGATGCGCT 240
DB 181 ATTCTGGGAATCAGGGCTCTTTTAATAAGGTCATCCAGCTGAATGATGCGCT 240
QY 241 GACTCAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 241 GACTCAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
DB 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
QY 361 CTAGGTTCCGATTGACCTGCAACTCTGACACCCACCTGCTTCAGGGGAGAGCTTGACC 420
DB 361 CTAGGTTCCGATTGACCTGCAACTCTGACACCCACCTGCTTCAGGGGAGAGCTTGACC 420
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGCT 480
DB 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGCT 480
QY 481 AAAAAATACAGAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGGATGTGC 540
DB 481 AAAAAATACAGAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGGATGTGC 540
QY 541 ACCTGGAATGCACTGTCTTTCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCTTG 600
DB 541 ACCTGGAATGCACTGTCTTTCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCTTG 600
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 601 GTGCTAGCTTCCAGAGGCTCCAG 626

```

RESULT 13  
 US-09-891-119A-8  
 ; Sequence 8, Application US/09891119A  
 ; Publication No. US20040013683A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Maddon, Paul J.  
 ; TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
 ; FILE REFERENCE: 24577-CY-B  
 ; CURRENT APPLICATION NUMBER: US/09/891,119A  
 ; CURRENT FILING DATE: 2001-06-25  
 ; NUMBER OF SEQ ID NOS: 22  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 8  
 ; LENGTH: 1742  
 ; TYPE: DNA

ORGANISM: Human  
 US-09-891-119A-8

Query Match 43.2%; Score 622.8; DB 11; Length 1742;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTGCTTGTGCTGCTCACTGGCGCTCCCTCCA 60
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTGCTTGTGCTGCTCACTGGCGCTCCCTCCA 135
QY 61 GCAGCCACTCAGAGGAAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACCTGACC 120
DB 136 GCAGCCACTCAGAGGAAAGAAAGTGTGCTGGCCAAAAGGGGATACAGTGAACCTGACC 195
QY 121 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCACCAATTAAG 180
DB 196 TGTACAGCTTCCCAAGAGAGCATATTCACCTGGAATACTCCACCAATTAAG 255
QY 181 ATTCTGGGAATCAGGGCTCTTTTAATAAGGTCATCCAGCTGAATGATGCGCT 240
DB 256 ATTCTGGGAATCAGGGCTCTTTTAATAAGGTCATCCAGCTGAATGATGCGCT 315
QY 241 GACTCAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAGAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
DB 376 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 435
QY 361 CTAGGTTCCGATTGACCTGCAACTCTGACACCCACCTGCTTCAGGGGAGAGCTTGACC 420
DB 436 CTAGGTTCCGATTGACCTGCAACTCTGACACCCACCTGCTTCAGGGGAGAGCTTGACC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGCT 480
DB 486 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGCT 555
QY 481 AAAAAATACAGAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGGATGTGC 540
DB 556 AAAAAATACAGAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGGATGTGC 615
QY 541 ACCTGGAATGCACTGTCTTTCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCTTG 600
DB 616 ACCTGGAATGCACTGTCTTTCAGAAACCAAGAAAGTGAAGTTCAAAATAGACATCTTG 675
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 676 GTGCTAGCTTCCAGAGGCTCCAG 701

```

RESULT 14  
 US-10-024-329-1  
 ; Sequence 1, Application US/10024329  
 ; Publication No. US20030157063A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: SANHADJI, Kamel  
 ; APPLICANT: TOURAINE, Jean-Louis  
 ; APPLICANT: LEROY, Pierre  
 ; APPLICANT: MEHTALI, Majid  
 ; TITLE OF INVENTION: Gene therapy using anti-gp41 antibody and cd4 immunoadhesin  
 ; FILE REFERENCE: 109993  
 ; CURRENT APPLICATION NUMBER: US/10/024,329  
 ; CURRENT FILING DATE: 2001-12-21  
 ; NUMBER OF SEQ ID NOS: 33  
 ; SOFTWARE: PatentIn Ver. 2.1  
 ; SEQ ID NO 1  
 ; LENGTH: 1745  
 ; TYPE: DNA  
 ; ORGANISM: human bcd4  
 US-10-024-329-1

Query Match 43.2%; Score 622.8; DB 15; Length 1745;  
 Best Local Similarity 99.7%; Pred. No. 1.4e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 60
DB 79 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 138
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120
DB 139 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 198
QY 121 TGTACAGCTTCCGAGAAAGACATACATTCACCTGGAATACTCCACAGATAAG 180
DB 199 TGTACAGCTTCCGAGAAAGACATACATTCACCTGGAATACTCCACAGATAAG 258
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 259 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 318
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300
DB 319 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 378
QY 301 ATGAAGACTCAGATTAATCTGTAAGTGAAGGACCAAGAGAGAGTGCATTTG 360
DB 379 ATGAAGACTCAGATTAATCTGTAAGTGAAGGACCAAGAGAGAGTGCATTTG 438
QY 361 CTAGTGTTCGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCGAGAGCTTAAC 420
DB 439 CTAGTGTTCGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCGAGAGCTTAAC 498
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGAATGTAGAGTCCAAGGGGT 480
DB 499 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGAATGTAGAGTCCAAGGGGT 558
QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540
DB 559 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 618
QY 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATATGACATGCTG 600
DB 619 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATATGACATGCTG 678
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 679 GTGCTAGCTTCCAGAGGCTCCAG 704

```

## RESULT 15

US-10-207-655-169  
 ; Sequence 169, Application US/10207655  
 ; Publication No. US20030118592A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Leebetter, Jeffrey A.  
 ; APPLICANT: Hayden-Leebetter, Martha S.  
 ; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS  
 ; FILE REFERENCE: 380069.401C1  
 ; CURRENT APPLICATION NUMBER: US/10/207,655  
 ; CURRENT FILING DATE: 2002-07-25  
 ; NUMBER OF SEQ ID NOS: 426  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 169  
 ; LENGTH: 3084  
 ; TYPE: DNA  
 ; ORGANISM: Homo sapiens  
 ; US-10-207-655-169

Query Match 43.2%; Score 622.8; DB 15; Length 3084;  
 Best Local Similarity 99.7%; Pred. No. 1.7e-170;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 60

```

DB 153 ATGAACCGGGAGTCCCTTTTAAAGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 212
QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 120
DB 213 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAATGACC 272
QY 121 TGTACAGCTTCCGAGAAAGACATACATTCACCTGGAATACTCCACAGATAAG 180
DB 273 TGTACAGCTTCCGAGAAAGACATACATTCACCTGGAATACTCCACAGATAAG 332
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 240
DB 333 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCCAGCTGAATGATCGGCT 392
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300
DB 393 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 452
QY 301 ATGAAGACTCAGATTAATCTGTAAGTGAAGGACCAAGAGAGAGTGCATTTG 360
DB 453 ATGAAGACTCAGATTAATCTGTAAGTGAAGGACCAAGAGAGAGTGCATTTG 512
QY 361 CTAGTGTTCGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCGAGAGCTTAAC 420
DB 513 CTAGTGTTCGATTGACTGCCAATCTGACACCCACTGCTTCAGGGGCGAGAGCTTAAC 572
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGAATGTAGAGTCCAAGGGGT 480
DB 573 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGAATGTAGAGTCCAAGGGGT 632
QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 540
DB 633 AAAAACAATACAGGGGGGAAAGACCTCTCCGTCTCAGCTGAGCTCCAGATATGAGC 692
QY 541 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATATGACATGCTG 600
DB 693 ACCTGACATGCACTGTCTTGCAAGACCAAGAAAGGTGAGTTCAAAATATGACATGCTG 752
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626
DB 753 GTGCTAGCTTCCAGAGGCTCCAG 778

```

Search completed: November 22, 2004, 16:11.13  
 Job time : 778 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 09:24:30 ; Search time 4897 Seconds  
(without alignments)  
10715.380 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440  
Sequence: 1 atgaacccgggagctccctt.....gtcagcagtcataactatg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 1821985908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : EST:

1: gb\_est1:\*  
2: gb\_est2:\*  
3: gb\_hic:\*  
4: gb\_est3:\*  
5: gb\_est4:\*  
6: gb\_est5:\*  
7: gb\_est6:\*  
8: gb\_g881:\*  
9: gb\_g882:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	622.8	43.2	783	6	CD609192 56069931J
2	622.8	43.2	932	5	BX438745 BX438745
3	610.2	42.4	816	4	BI838409 603083373
4	599.4	41.6	822	5	BX457159 BX457159
5	598.8	41.6	801	4	BI822118 603039813
6	589	40.9	840	4	BI838893 603087279
7	577.4	40.1	791	4	BI819727 603041415
8	539	37.4	713	5	BX437619 BX437619
9	519.4	36.1	723	4	BI915265 603184313
10	506.4	35.2	679	6	CD609191 56069931H
11	504	35.0	732	1	AU141298 AU141298
12	488.8	33.9	754	7	CO246446 AGENCOURT
13	488.4	33.9	710	6	CF125444 UI-HF-EL0
14	483.8	33.6	694	7	CB052686 NISC_g109
15	480.6	33.4	803	6	CO249540 AGENCOURT
16	466	32.4	809	7	CO248894 AGENCOURT
17	457.4	31.8	748	5	BX378811 BX378811
18	452.8	31.4	821	7	CO245985 AGENCOURT
19	429.4	28.8	616	6	CF125140 UI-HF-EL0
20	407.6	28.3	751	7	CO248544 AGENCOURT
21	402.6	28.0	687	2	AW752367 RC2-CT020
22	374.6	25.0	767	1	AU136822 AU136822
23	366.6	25.5	686	6	CB052685 NISC_g109
24	364.2	25.3	755	7	CO246525 AGENCOURT

C 25	354.8	24.6	638	2	AW752358
26	345	24.0	893	1	AL553528
27	328.8	22.8	707	7	CK835626
28	302.8	21.0	567	6	CB997693 AGENCOURT
29	261.2	18.1	336	2	AM407282 UI-HF-BL0
30	261.2	18.1	374	2	AM406380 UI-HF-BL0
31	261.2	18.1	733	4	BG685053 602636816
32	259.6	18.0	325	2	AM407857 UI-HF-BL0
33	259.6	18.0	409	2	AM406057 UI-HF-BL0
34	259.6	18.0	429	6	CD707673 EST24200
35	259.6	18.0	448	2	AM404726 UI-HF-BL0
36	259.6	18.0	499	6	CD703936 EST20463
37	259.6	18.0	514	2	AM405207 UI-HF-BL0
38	259.6	18.0	643	4	BM782582 K-EST0059
39	259.6	18.0	723	4	BM007674 603617103
40	259.6	18.0	1047	5	BU146881 AGENCOURT
41	258	17.9	330	2	AM406212 UI-HF-BL0
42	258	17.9	414	6	CD693211 EST9734 h
43	258	17.9	425	2	AM406785 UI-HF-BL0
44	258	17.9	465	2	AM404697 UI-HF-BL0
45	258	17.9	471	6	CD687139 EST3660 h

## ALIGNMENTS

```

RESULT 1
LOCUS      CD609192
DEFINITION 56069931J1 FLP Homo sapiens cDNA, mRNA sequence.
ACCESSION  CD609192
VERSION    CD609192.1 GI:40257455
KEYWORDS   EST.
SOURCE     Homo sapiens
ORGANISM   Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE  1 (bases 1 to 783)
AUTHORS   Fu,G.K., Wang,J.T., Yang,J., Au-Young,J. and Stuve,L.L.
TITLE     Circular rapid amplification of cDNA ends for high-throughput
          extension cloning of partial genes
JOURNAL   Genomics 84 (1), 205-210 (2004)
COMMENT   Contact: Fu GK
          Incyte Genomics, Inc.
          3160 Porter Dr., Palo Alto, CA 94304, USA
          Tel: 6508454102
          Email: gfu@incyte.com.

FEATURES             location/Qualifiers
     source            1..783
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone_lib="FLP"
                     /note="Vector: pDrive Cloning Vector"

ORIGIN
Query Match      43.2%; Score 622.8; DB 6; Length 783;
Best Local Similarity 99.7%; Pred. No. 2.4e-152;
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGCTGCTGCAACTGCGCTCTCCCA 60
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
726 ATGAACCGGGAGTCCCTTTTAGGCACTTCTTGCTGCTGCAACTGCGCTCTCCCA 667
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
61 GAAGCACTCCAGGAGAAAGTGTGCTGGGCAAAAAGGAGATACAGTGAACCTGACC 120
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
666 GAAGCACTCCAGGAGAAAGTGTGCTGGGCAAAAAGGAGATACAGTGAACCTGACC 607
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
121 TGTACAGCTTCCAGAGAAAGCATACATTCACCTGGAATACTCCACAGATAAAG 180
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
606 TGTACAGCTTCCAGAGAAAGCATACATTCACCTGGAATACTCCACAGATAAAG 547
| | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
181 ATTGGGAATCAGGGCTCTTTTAACTAAAGTCCATCCAGCTGAATGATGGCGCT 240

```

```

|||||
Db 546 ATCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCATCAAGTGAATGATCGGCT 487
Qy 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAACCTCCCGATCATCAAGAACTTTAG 300
Db 486 GACTCAAGAAAGAGCTTTGGGACCAAGAAACCTCCCGATCATCAAGAACTTTAG 427
Qy 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 360
Db 426 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTCAATTG 367
Qy 361 CTAGTGTTCGATTTGACTGCTGCAACCTGACACCCAGCTTCCAGGGGCAAGGCTGACC 420
Db 366 CTAGTGTTCGATTTGACTGCTGCAACCTGACACCCAGCTTCCAGGGGCAAGGCTGACC 307
Qy 421 CTGACCTTGGAGAGCCCCCTGATAGACCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 480
Db 306 CTGACCTTGGAGAGCCCCCTGATAGACCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 247
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATATGTGC 540
Db 246 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATATGTGC 187
Qy 541 ACCTGGAATGCACTGTCTTGGACAGAACCAAGAAAGTGGAGTTCAAAATATGACATGTG 600
Db 186 ACCTGGAATGCACTGTCTTGGACAGAACCAAGAAAGTGGAGTTCAAAATATGACATGTG 127
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db 126 GTGCTAGCTTTCCAGAAAGGCTCCAG 101

```

```

RESULT 2
LOCUS BX438745 932 bp mRNA linear EST 05-MAY-2004
DEFINITION BX438745 Homo sapiens PLACENTA Homo sapiens cDNA clone CS0DE002YG01
ACCESSION BX438745
VERSION BX438745.2 GI:47035651
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 932)
L.J.W.B., Gruber, C., Jessee, J. and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
On May 15, 2003 this sequence version replaced gi:30787731.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 Evry cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library
was not normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
This sequence belongs to sequence cluster 6485.r
For more information about this cluster, see
http://www.genoscope.cns.fr/cdna?b=CS0DE002AD01QP1&c=6485.r.
Location/Qualifiers
1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YG01"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6, 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

## FEATURES

source

```

1..932
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DE002YG01"
/tissue_type="PLACENTA"
/clone_lib="Homo sapiens PLACENTA"
/notes="Vector: pCMVSPORT 6, 1st strand cDNA was primed
with a NotI-oligo (dT) primer. Five prime end enriched,
double-strand cDNA was digested with Not I and cloned into
the Not I and EcoRV sites of the pCMVSPORT 6 vector."

```

Library was not normalized."

Query Match 43.2%; Score 622.8; DB 5; Length 932;  
 Best Local Similarity 99.7%; Pred. No. 2,5e-152;  
 Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```

Qy 1 ATGAAACGGGAGATCCCTTTTAAAGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 60
Db 180 ATGAAACGGGAGATCCCTTTTAAAGCACTTCTGTGTGTGCAACTGGCGCTCTCCCA 239
Qy 61 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACATGTAACCTGACC 120
Db 240 GCAGCCACTCAGGAGAAAGAGTGTCTGGGCAAAAAGGGATACATGTAACCTGACC 299
Qy 121 TGTAGAGCTTCCAGAAAGAGATATCAATTCACCTGGAATAAATCCATCAAGATTAAG 180
Db 300 TGTAGAGCTTCCAGAAAGAGATATCAATTCACCTGGAATAAATCCATCAAGATTAAG 359
Qy 181 ATCTGGGAATACAGGGCTCTCTTCTTAATTAAGTTCATCAAGCTGAATGATCGGCT 240
Db 360 ATCTGGGAATACAGGGCTCTCTTCTTAATTAAGTTCATCAAGCTGAATGATCGGCT 419
Qy 241 GACTCAAGAAAGAGCTTTGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAG 300
Db 420 GACTCAAGAAAGAGCTTTGGACCAAGAAACTTCCCTGATCATCAAGAACTTTAG 479
Qy 301 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 360
Db 480 ATGAAAGACTCAGATTAATCATCTGTGAAGTGAGGACCAAGAGAGAGTGAATTG 539
Qy 361 CTAGTGTTCGATTTGACTGCTGCAACCTGACACCCAGCTTCCAGGGGCAAGCTGACC 420
Db 540 CTAGTGTTCGATTTGACTGCTGCAACCTGACACCCAGCTTCCAGGGGCAAGCTGACC 599
Qy 421 CTGACCTTGGAGAGCCCCCTGATAGACCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 480
Db 600 CTGACCTTGGAGAGCCCCCTGATAGACCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 659
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATATGTGC 540
Db 660 AAAAACAATACAGGGGGGGAAGACCTCTCGTGTCTGAGCTGAGCTCCAGATATGTGC 719
Qy 541 ACCTGGAATGCACTGTCTTGGACAGAACCAAGAAAGTGGAGTTCAAAATATGACATGTG 600
Db 720 ACCTGGAATGCACTGTCTTGGACAGAACCAAGAAAGTGGAGTTCAAAATATGACATGTG 779
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db 780 GTGCTAGCTTTCCAGAAAGGCTCCAG 805

```

```

RESULT 3
LOCUS BI838409 816 bp mRNA linear EST 04-OCT-2001
DEFINITION BI838409 Homo sapiens cDNA clone IMAGE:5222323 5',
ACCESSION BI838409
VERSION BI838409.1 GI:15949959
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 816)
NIH-MGC http://mgc.ncl.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: sgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

```

```

REFERENCE
AUTHORS NIH-MGC
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: sgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LIML)

```

DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/BLNL at:  
<http://image.llnl.gov>

Plate: L1M1559 row: b column: 20  
 High quality sequence stop: 803.  
 Location/Qualifiers

## FEATURES

source

1. 816  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:522323"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector:  
 pCMV-Sport6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
 source anonymous pool of spleen and pancreas from 28 yo  
 male. Library is oligo-dT primed and directionally cloned  
 (EcoRV site is destroyed upon cloning). Average insert  
 size 1.5 kb, insert size range 1-2.5 kb. Library is  
 normalized and enriched for full-length clones and was  
 constructed by C. Gruber (Invitrogen). Research Genetics  
 tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

Query Match 42.4%; Score 610.2; DB 4; Length 816;  
 Best Local Similarity 99.4%; Pred. No. 5e-149;  
 Matches 623; Conservative 0; Mismatches 3; Indels 1; Gaps 1;

```

Oy 1 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGTCACATGCGCGCTCTCCCA 60
Db 167 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGTCACATGCGCGCTCTCCCA 226
Oy 61 GCAGCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACACC 120
Db 227 GCAGCACTCAGGAGAAAGAAAGTGTGCTGGGCAAAAAAGGGATACAGTGAACACC 286
Oy 121 TGTACAGCTTCCCAAGAAAGAGCATACATTCCTGGAATACTCCACCAAGTAAG 180
Db 287 TGTACAGCTTCCCAAGAAAGAGCATACATTCCTGGAATACTCCACCAAGTAAG 346
Oy 181 ATTCTGGGAATCAGAGGCTCTTCTTAATTAAGGTCATCCAGCTGAATGATGCGCT 240
Db 347 ATTCTGGGAATCAGAGGCTCTTCTTAATTAAGGTCATCCAGCTGAATGATGCGCT 406
Oy 241 GACTCAGAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300
Db 407 GACTCAGAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 466
Oy 301 ATAGAAGCTCAGATATTCTTACATCTGTGAGTGAAGACCAAGAGAGAGGTGCAATTG 360
Db 467 ATAGAAGCTCAGATATTCTTACATCTGTGAGTGAAGACCAAGAGAGAGGTGCAATTG 526
Oy 361 CTAGGTTCCGATGACCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420
Db 527 CTAGGTTCCGATGACCTGCAACTCTGACACCCACTGCTTCAAGGGGAGAGCTGACC 586
Oy 421 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCAGTGCATGTAGAGATCCAGGGGT 480
Db 587 CTGACCTTGGAGAGCCCTCTGTGTAGTACCCCTCAGTGCATGTAGAGATCCAGGGGT 646
Oy 481 AAAAAATACAGAGGGGGGAAAGCCTCTTCC-GTGTCTCAGCTGAGCTCCAGGAATGTGG 539
Db 647 AAAAAATACAGAGGGGGGAAAGCCTCTTCCGTGTCTCAGCTGAGCTCCAGGAATGTGG 706
Oy 540 CACCTGACATGCACTGTCTTGACAGACCAAGAAAGGTGAGTCAAAATAGATGATGT 599
Db 707 CACCTGACATGCACTGTCTTGACAGACCAAGAAAGGTGAGTCAAAATAGATGATGT 766
Oy 600 GGTGCTAGCTTTCCAGAGGCTCTCCG 626
Db 767 GGTGCTAGCTTTCCAGAGGCTCTCCG 793

```

## RESULT 4

BX457159

LOCUS

DEFINITION

ACCESSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

BX457159 822 bp mRNA linear EST 05-MAY-2004  
 BX457159 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP006YP07  
 5-PRIME, mRNA sequence.  
 BX457159  
 BX457159.2 GI:47051693  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 1 (bases 1 to 822)  
 Li, W.B., Gruber, C., Jessee, J., and Polayes, D.  
 Full-length cDNA libraries and normalization  
 Unpublished (2001)  
 On May 22, 2003 this sequence version replaced gi:31020883.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: [seqref@genoscope.cns.fr](mailto:seqref@genoscope.cns.fr), Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library  
 was not normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see  
<http://www.genoscope.cns.fr/cdnas7-CS0CAP006CH04QP1&c=6485.r>.

## FEATURES

source

1. 822  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP006YP07"  
 /tissue\_type="THYMUS"  
 /clone\_lib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed  
 with a NotI-oligo(dT) primer. Five prime end enriched,  
 double-strand cDNA was digested with Not I and cloned into  
 the Not I and EcoRV sites of the pCMVSPORT 6 vector.  
 Library was not normalized."

## ORIGIN

Query Match 41.6%; Score 599.4; DB 5; Length 822;  
 Best Local Similarity 99.2%; Pred. No. 3.4e-146;  
 Matches 622; Conservative 1; Mismatches 2; Indels 2; Gaps 2;

```

Oy 1 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGTCACATGCGCGCTCTCCCA 60
Db 183 ATGAACCGGGAGTCCCTTTTATGACCTTGTGTCGTCACATGCGCGCTCTCCCA 242
Oy 61 G-CAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTAC 119
Db 243 GSCAGCCACTCAGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACCTAC 302
Oy 120 CTGTACAGCTTCCCAAGAAAGAGCATATTCCTCAGTGAATACTCAACCAAGATTA 179
Db 303 CTGTACAGCTTCCCAAGAAAGAGCATATTCCTCAGTGAATACTCAACCAAGATTA 362
Oy 180 GATTCTGGGAATACAGAGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATGCGC 239
Db 363 GATTCTGGGAATACAGAGGCTCTTCTTAATTAAGTCCATCAAGCTGAATGATGCGC 422
Oy 240 TGACCTCAAGAAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAACTTA 299
Db 423 TGACCTCAAGAAAGAGCCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAACTTA 482
Oy 300 GATAGAAGACTCAGATATTCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATT 359
Db 483 GATAGAAGACTCAGATATTCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATT 542

```



```

source
1. .840
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5226427"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_120"
/notes="Organ: pooled pancreas and spleen; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of spleen and pancreas from 28 yo
male. Library is oligo-dT primed and directionally cloned
(EcoRV site is destroyed upon cloning). Average insert
size 1.5 kb, insert size range 1-2.5 kb. Library is
normalized and enriched for full-length clones and was
constructed by C. Gruber (Invitrogen). Research Genetics
tracking code 025. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      40.9%; Score 589; DB 4; Length 840;
Best Local Similarity 98.1%; Pred. No. 1.9e-143;
Matches 617; Conservative 0; Mismatches 10; Indels 2; Gaps 2;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCACTGCGCTCTCCCA 60
DB 168 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCACTGCGCTCTCCCA 227
QY 61 GGAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAATGAC 120
DB 228 GGAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAATGAC 287
QY 121 TGTACAGCTCCAGAAAGAGCATACATTCCTGAAAAAATCCAAACGATTAAG 180
DB 288 TGTACAGCTCCAGAAAGAGCATACATTCCTGAAAAAATCCAAACGATTAAG 347
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCCAGCTGATGCGCT 240
DB 348 ATTCTGGGAATCAGGGCTCTTCTTAATTAAGGTCATCCAGCTGATGCGCT 407
QY 241 GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATTTAAG 300
DB 408 GACTCAAGAAAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATTTAAG 467
QY 301 ATAGAAGACTGATATACATCTGTGAAGTGGAGGACCAAGAGGAGGTCAATTG 360
DB 468 ATAGAAGACTGATATACATCTGTGAAGTGGAGGACCAAGAGGAGGTCAATTG 527
QY 361 CTAGTGTTCGATGACTGCAACTCTGACCACTGCTTCAGGGGACAGAGCTGACC 420
DB 528 CTAGTGTTCGATGACTGCAACTCTGACCACTGCTTCAGGGGACAGAGCTGACC 587
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTGAGAGGT 480
DB 588 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTCAATGTGAGAGGT 647
QY 481 AAA-AACATACAGGGGGGAGAACCTCTCCGTGTCTCAGCTGAGGCTCAAGATATG 539
DB 648 AAAACAACATACAGGGGGGAGAACCTCTCCGTGTCTCAGCTGAGGCTCAAGATATG 707
QY 540 CACCTGGAATGCACTGTCTTCAGAAACCAAGAAAGTGTGAGTTCAAAATGACATCT 599
DB 708 CACCTGGAATGCACTGTCTTCAGAAACCAAGAAAGTGTGAGTTCAAAATGACATCT 767
QY 600 GG-TGCTAGCTTTCCAGAAAGGCTCCGGA 627
DB 768 GGAGGCTAGCTTTCCAGAAAGGCTCCGGA 796

```

```

RESULT 7
BI819727      791 bp      mRNA      linear      EST 04-OCT-2001
LOCUS        603041415p1 NIH_MGC_115 Homo sapiens cDNA clone IMAGE:5182274 5',
DEFINITION   mRNA sequence.
ACCESSION    BI819727

```

```

VERSION      BI819727.1 GI:15931277
KEYWORDS     EST.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    NIH-MGC http://mgs.mci.nih.gov/.
AUTHORS      National Institutes of Health, Mammalian Gene Collection (MGC)
TITLE        Unpublished (1999)
JOURNAL      Contact: Robert Strausberg, Ph.D.
COMMENT      Email: cgaabs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
              DNA Sequencing by: Incyte Genomics, Inc.
              Clone distribution: MGC clone distribution information can be
              found through the I.M.A.G.E. Consortium/LNL at:
              http://image.llnl.gov
              Plate: LLM11454 row: n column: 03
              High quality sequence stop: 782.

FEATURES
source
1. .791
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5182274"
/lab_host="DH10B"
/clone_1ib="NIH_MGC_115"
/notes="Organ: pooled brain, lung, testis; Vector:
pcmv-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA
source anonymous pool of 6 male brains, age range 23-27; 1
male lung, age 27; and 1 male testis, age 69. Library is
oligo-dT primed and directionally cloned (EcoRV site is
destroyed upon cloning). Average insert size 1.8 kb,
insert size range 1-3 kb. Library is normalized and
enriched for full-length clones and was constructed by C.
Gruber (Invitrogen). Research Genetics tracking code
021. Note: this is a NIH_MGC Library."

ORIGIN
Query Match      40.1%; Score 577.4; DB 4; Length 791;
Best Local Similarity 99.3%; Pred. No. 2.1e-140;
Matches 611; Conservative 0; Mismatches 1; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGCGCTCTCCCA 60
DB 176 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGCGCTCTCCCA 235
QY 61 GGAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAATGAC 120
DB 236 GGAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGATACAGTGAATGAC 295
QY 121 TGTACAGCTCCAGAAAGAGCATACATTCCTGAAAAAATCCAAACGATTA-A 179
DB 236 TGTACAGCTCCAGAAAGAGCATACATTCCTGAAAAAATCCAAACGATTA-A 355
QY 180 GATTCGGGAATCAGGGCTCTTCTTAATAAGTCCATCAAGCTGAATGATGCGCG 239
DB 236 TGTACAGCTCCAGAAAGAGCATACATTCCTGAAAAAATCCAAACGATTA-A 355
QY 356 GATTCGGGAATCAGGGCTCTTCTTAATAAGTCCATCAAGCTGAATGATGCGCG 415
DB 240 TGACTCAAGAAAGAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAA 299
QY 416 TGACTCAAGAAAGAGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAA 475
DB 300 GATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCAATT 359
QY 476 GATAGAAGACTCAGATACCTTACATCTGTGAAGTGAAGGACCAAGAGGAGGTCAATT 535
QY 360 GCTAGTGTTCGATTTGATGCTGCAACTGTGACACCACTGCTTTCAGGGGACAGAGCTTAC 419
DB 536 GCTAGTGTTCGATTTGATGCTGCAACTGTGACACCACTGCTTTCAGGGGACAGAGCTTAC 595

```

Qy 420 CTTGACCTTGG-AGAGCCCCCTGTAGTAGAGCCCTCAGTCAATGTAGAGTCCAAAGG 478  
 Db 596 CTTGACCTTGGCAAGCCCCCTGTAGTAGAGCCCTCAGTCAATGTAGAGTCCAAAGG 655  
 Qy 479 GTAAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATATAGT 538  
 Db 656 GTAAAAACATACAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATATAGT 715  
 Qy 539 GCACCTGAGAC-TCGACCTGTCTTGCAAGAACGAAAGAGTGGAGTCAAAATAGACATC 597  
 Db 716 GCACCTGAGACCTGTCTGTGCAAGAACGAAAGAGTGGAGTCAAAATAGACATC 775  
 Qy 598 GTGCTGCTAGCTTC 612  
 Db 776 GTGCTGCTAGCTTC 790

RESULT 8  
 BX437619 713 bp mRNA linear EST 04-MAY-2004  
 LOCUS BX437619 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP007YD02  
 DEFINITION BX437619 Homo sapiens THYMUS Homo sapiens cDNA clone CS0CAP007YD02  
 ACCESSION BX437619  
 VERSION BX437619.2 GI:46999918  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 713)  
 AUTHORS Li, W.-B., Gruber, C., Jessee, J., and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On May 15, 2003 this sequence version replaced gi:30771588.  
 CONTACT: Genoscope  
 Genoscope - Centre National de Sequencage  
 BP 191 91006 Evry cedex - France  
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized. Library was constructed by Life Technologies, a division of Invitrogen.  
 This sequence belongs to sequence cluster 6485.r  
 For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdnats-CS0CAP007YD01Q1&c=6485.r.  
 FEATURES  
 source  
 1..713  
 Location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0CAP007YD02"  
 /cbase="THYMUS"  
 /clone\_lib="Homo sapiens THYMUS"  
 /note="Vector: pCMVSPORT 6; 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoRV sites of the pCMVSPORT 6 vector. Library was not normalized."  
 ORIGIN  
 Query Match 37.4%; Score 539; DB 5; Length 713;  
 Best Local Similarity 99.5%; Pred. No. 2.5e-110; Indels 1; Gaps 1;  
 Matches 550; Conservative 1; Mismatches 1;  
 Qy 1 ATGAACGGGAGTCCCTTTAGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 60  
 Db 161 ATGAACGGGAGTCCCTTTAGCACTTGTCTGTGTGCAACTGGCGCTCTCCCA 220  
 Qy 61 GCAGCCACTCAGGAAAGAAAGTGTGTGGCAAAAAGGGATACAGTGAATGACC 120  
 Db 221 GCAGCCACTCAGGAAAGAAAGTGTGTGGCAAAAAGGGATACAGTGAATGACC 280

Qy 121 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAAATCCCAACGATTAAG 180  
 Db 281 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAAATCCCAACGATTAAG 340  
 Qy 181 ATTCTGGAAATCAGGGCTCTCTTTACTTAAGATTCATCCAAAGTGAATGATCGGCT 240  
 Db 341 ATTCTGGAAATCAGGGCTCTCTTTACTTAAGATTCATCCAAAGTGAATGATCGGCT 400  
 Qy 241 GACTCAGAAAGAAAGCTTTGGGACCAAGAAAGTCCCTGATCATCAAGAAATCTTAAG 300  
 Db 401 GACTCAGAAAGAAAGCTTTGGGACCAAGAAAGTCCCTGATCATCAAGAAATCTTAAG 460  
 Qy 301 ATGAAGACTCAGATTAATCACTGTGAGAGTGGAGACCAAGAGAGAGTGCATTTG 360  
 Db 461 ATGAAGACTCAGATTAATCACTGTGAGAGTGGAGACCAAGAGAGAGTGCATTTG 520  
 Qy 361 CTAGTGTTCGAGATTGACTGCCAACTGTGACACCACTGCTTCAGGGGCAAGGCTTACC 420  
 Db 521 CTAGTGTTCGAGATTGACTGCCAACTGTGACACCACTGCTTCAGGGGCAAGGCTTACC 580  
 Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGCAGATGTAGAGTCCAAAGGGGT 480  
 Db 581 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTCAGTGCAGATGTAGAGTCCAAAGGGGT 640  
 Qy 481 AAAAATACATCA-GGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATTAAGTGG 539  
 Db 641 AAAAATACATCAAGGGGGGGAAGACCTCTCCGTGTCTCAGCTGAGAGTCCAGATTAAGTGG 700  
 Qy 540 CACCTGACATGC 552  
 Db 701 CACCTGACATGC 713

RESULT 9  
 B1915265 723 bp mRNA linear EST 16-OCT-2001  
 LOCUS B1915265  
 DEFINITION B1915265 NIH\_MGC\_121 Homo sapiens cDNA clone IMAGE:5248226 5', mRNA sequence.  
 ACCESSION B1915265  
 VERSION B1915265.1 GI:16179363  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens (human)  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 723)  
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strauberg, Ph.D.  
 Email: sgabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 cDNA Library Preparation: Life Technologies, Inc.  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LNL11626 row: J column: 03  
 High quality sequence stop: 717.  
 Location/Qualifiers  
 1..723  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5248226"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH MGC 121"  
 /note="Organ: brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 3 fetal brains, female age 20 weeks, female age 24 weeks, and male age 26 weeks. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon



Fax: 81-438-52-3986  
 Email: [genomics@hri.co.jp](mailto:genomics@hri.co.jp)  
 HRI human cDNA project; 5'- & 3'-end one pass sequencing: Helix  
 Research Institute; cDNA library construction: Department of  
 Virology, Institute of Medical Science, University of Tokyo, and  
 Helix Research Institute.

# FEATURES

Source  
 Location/Qualifiers  
 1. 732  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="THYROID0359"  
 /tissue="thyroid gland"  
 /clone\_id="THYROID"  
 /note="Vector: pME18FLJ3"

## ORIGIN

Query Match 35.0%; Score 504; DB 1; Length 732;  
 Best Local Similarity 97.2%; Pred. No. 4e-121;  
 Matches 52; Conservative 0; Mismatches 13; Indels 2; Gaps 1;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCACTGGGCTCTCCCA 60
   |||||||
DB 196 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCACTGGGCTCTCCCA 255
   |||||||
QY 61 GCAGCCACTAGGGAAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 120
   |||||||
DB 256 GCAGCCACTAGGGAAGAAAGTGTGCTGGCAAAAAGGGATAGAGTGAATGACC 315
   |||||||
QY 121 TGTACAGCTTCCGAAAGAAAGCATACATTCACCTGGAATACTCCACAGATTAAG 180
   |||||||
DB 316 TGTACAGCTTCCGAAAGAAAGCATACATTCACCTGGAATACTCCACAGATTAAG 375
   |||||||
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 240
   |||||||
DB 376 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 435
   |||||||
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
   |||||||
DB 436 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 495
   |||||||
QY 301 ATGAAGACTCAGTACTTACTCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 360
   |||||||
DB 496 ATGAAGACTCAGTACTTACTCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 555
   |||||||
QY 361 CTAGTGTTCGATGACTGCACTGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 420
   |||||||
DB 556 CTAGTGTTCGATGACTGCACTGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 615
   |||||||
QY 421 CTGACCTTGAAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGTAAGAGTCAAGGGGT 480
   |||||||
DB 616 CTGACCTTGAAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGTAAGAGTCAAGGGGT 675
   |||||||
QY 481 AAAAACAATACAGGGGGG--AAGACCTCTCTGCTTCACTGAGCTCCAGATGATA 535
   |||||||
DB 676 AAAAACAATACAGGGGGGAAAGACCTTTTCCGGGTCTAATGAGACTTCANAGATA 732
   |||||||

```

RESULT 12  
 CO246446 754 bp mRNA linear EST 23-JUN-2004

LOCUS  
 DEFINITION  
 AGENCOURT 26517764 NIH MGC 212 Homo sapiens cDNA clone  
 IMAGE:30924065 5', mRNA sequence.

ACCESSION  
 CO246446  
 CO246446.1 GI:49111174

VERSION  
 KEYWORDS  
 EST.

ORGANISM  
 SOURCE  
 Homo sapiens (human)

REFERENCE  
 AUTHORS  
 1 (bases 1 to 754)  
 NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE  
 National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL

Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: [cgabds-remail.nih.gov](mailto:cgabds-remail.nih.gov)  
 Tissue Procurement: Mary Hendrix

cDNA Library Preparation: M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
 Plate: NDAM165 row: o column: 18  
 High quality sequence stop: 507.

## FEATURES

### Source

Location/Qualifiers

1. 754

/organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30924065"  
 /tissue\_type="Chondrosarcoma lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_id="NIH\_MGC\_212"  
 /note="Organ: Lung; Vector: pYX-Ase; Site 1: Ecor I;  
 Site 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction. Ligated  
 with Ecor I adaptor, digested with Not I and then cloned  
 directionally into pYX-Ase vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGCCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.9%; Score 488.8; DB 7; Length 754;  
 Best Local Similarity 96.6%; Pred. No. 4.1e-117;  
 Matches 510; Conservative 0; Mismatches 17; Indels 1; Gaps 1;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCACTGGGCTCTCCCA 60
   |||||||
DB 173 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCACTGGGCTCTCCCA 232
   |||||||
QY 61 GCAGCCACTAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATGACC 120
   |||||||
DB 233 GCAGCCACTAGGGAAGAAAGTGTGCTGGCAAAAAGGGATACAGTGAATGACC 292
   |||||||
QY 121 TGTACAGCTTCCGAAAGAAAGCATACATTCACCTGGAATACTCCACAGATTAAG 180
   |||||||
DB 293 TGTACAGCTTCCGAAAGAAAGCATACATTCACCTGGAATACTCCACAGATTAAG 352
   |||||||
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 240
   |||||||
DB 353 ATTCTGGGAATCAGGGCTCTTCTTAACCTAAAGTCCATCAAGTGAATGATCGGCT 412
   |||||||
QY 241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
   |||||||
DB 413 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 472
   |||||||
QY 301 ATGAAGACTCAGTACTTACTCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 360
   |||||||
DB 473 ATGAAGACTCAGTACTTACTCTGTGAAGTGAAGCAAGAAAGAGGTGCAATTG 532
   |||||||
QY 361 CTAGTGTTCGATGACTGCACTGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 420
   |||||||
DB 533 CTAGTGTTCGATGACTGCACTGCACTGCAACCACTGCTTCAAGGGGCAAGCTGACC 592
   |||||||
QY 421 CTGACCTTGAAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGTAAGAGTCAAGGGGT 480
   |||||||
DB 593 CTGACCTTGAAGAGCCCCCTGTGATAGCCCCCTCACTGCAATGTAAGAGTCAAGGGGT 652
   |||||||

```

QY 481 AAAACATACAGGGGGAGAACCTCTCGTGTCTACGTGAGCTC 528  
 DB 653 AAAA-ATACATGTAGGGGAAAACCTCTCTGTCTCATCTGATCTC 699

RESULT 13  
 CFI25444  
 LOCUS CFI25444  
 DEFINITION CFI25444 710 bp mRNA linear EST 05-AUG-2003  
 IMAGE:30563734 5', mRNA sequence.  
 CFI25444  
 CFI25444.1 GI:33201679  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 710)  
 AUTHORS Ronaldo,M.F., Lennon,G. and Soares,M.B.  
 TITLE Normalization and subtraction: two approaches to facilitate gene  
 discovery  
 JOURNAL Genome Res. 6 (9), 791-806 (1996)  
 MEDLINE 97044477  
 PUBMED 8889548  
 COMMENT Contact: Soares, MB  
 Coordinated Laboratory for Computational Genomics  
 University of Iowa  
 375 Newton Road, 4156 MEBRF, Iowa City, IA 52242, USA  
 Tel: 319 335 8250  
 Fax: 319 335 9565  
 Email: bento-soares@uiowa.edu  
 Tissue Procurement: Mary Hendrix  
 cDNA Library preparation: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Distribution information can be found at  
 http://genome.uiowa.edu/distribution/humanfl.html  
 Seq primer: pYX-5

FEATURES  
 source location/Qualifiers  
 1..710  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30563734"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1lb="N1H MGC 212"  
 /note="Organ: Lung; Vector: pYX-Asc; Site 1: EcoR I;  
 Site 2: Not I; The library was constructed according  
 Bonaldi, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with EcoR I adaptor, digested with Not I and then cloned  
 directionally into pYX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGGCCA. Tissue was provided by Mary Hendrix."

ORIGIN  
 Query Match 33.9%; Score 488.4; DB 6; Length 710;  
 Best Local Similarity 98.9%; Pred. No. 5.1e-117;  
 Matches 522; Conservative 0; Mismatches 3; Indels 3; Gaps 3;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACGTGGCGCTCTCCCA 60  
 DB 186 ATGAACCGGGAGTCCCTTTAGGCACTTCTGTGCTGCAACGTGGCGCTCTCCCA 245  
 QY 61 GCAGCACTCAGGAGAAAGAAAGTGTCTGTGGCAAAAAAGGGATACAGTGAATGACC 120  
 DB 246 GCAGCACTCAGGAGAAAGAAAGTGTGTGTGGCAAAAAAGGGATACAGTGAATGACC 305  
 QY 121 TGTACAGCTTCCAGAGAGACATACATTCACCTGGAAGAAAGTCCACACAGATAAG 180

DB 306 TGTACAGCTTCCAGAGAGAGCATACATTCACCTGGAAGAAAGTCCACAGATAAG 365  
 QY 181 ATTCTGGGAATACAGGCTCTCTTTAACTAAAGGTCATCCAGCTGAATATGCGCT 240  
 DB 366 ATTCTGGGAATACAGGCTCTCTTTAACTAAAGGTCATCCAGCTGAATATGCGCT 425  
 QY 241 GACTGAAGAAAGGCTTTGGAGCAAGAACTCCCTGATATATCAAGAAATCTTAAG 300  
 DB 426 GACTGAAGAAAGGCTTTGGAGCAAGAACTTTCCCTGATATATCAAGAAATCTTAAG 485  
 QY 301 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGAATTG 360  
 DB 486 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACCAAGAGAGAGGTGAATTG 545  
 QY 361 CTAGTGTTCGATGACTCCCACTCTGACACCCACTGCTTCAGGGGAGAGCTGACC 420  
 DB 546 CTAGTGTTCGATGACTCCCACTCTGACACCCACTGCTTCAGGGGAGAGCTGACC 605  
 QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTGATGCAATGTAGAGTCCAGGGGT 480  
 DB 606 CTGACCTTGGAGAG-CCCCCTGTGTAGTAC-CCCTGATGCAATGTAGAGTCCAGGGGT 663  
 QY 481 AAAACATACAGGGGGAGAACCTCTCGTGTCTACGTGAGCTC 528  
 DB 664 AAAACATACAGGGGGAGAAC-CTCTCGTGTCTACGTGAGCTC 710

RESULT 14  
 CB052686  
 LOCUS CB052686  
 DEFINITION NISC\_g109606.y1 NCI\_CGAP\_Le12 Homo sapiens cDNA clone IMAGE:3290578  
 5', mRNA sequence.  
 CB052686  
 CB052686.1 GI:27790973  
 EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 REFERENCE 1 (bases 1 to 694)  
 AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.  
 TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 JOURNAL Unpublished (1997)  
 COMMENT Contact: Robert Strausberg, Ph.D.  
 Email: cga@bbs-rcmail.nih.gov  
 cDNA Library preparation:  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium/LNL  
 Sequencing Center (NISC)  
 Clone Distribution: NCI-CGAP clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 info@image.lnl.gov  
 Plate: LLM8058 row: I column: 11  
 Seq primer: M13RP1 reverse primer (ABI).  
 Location/Qualifiers  
 1..694  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:3290578"  
 /tissue\_type="leiomyosarcoma"  
 /lab\_host="DH10B"  
 /clone\_1lb="NCI CGAP Le12"  
 /note="Organ: soft tissue; Vector: pRTD-Pac (Pharmacia)  
 with a modified polylinker; Site 1: Not I; Site 2: Eco RI;  
 1st strand cDNA was primed with a Not I - oligo(dT) primer  
 [5'-ACCTGGAAGAAATTCGGCGGCAATCTTTTATTTT-3'],  
 double-stranded cDNA was ligated to Eco RI adaptors  
 (Pharmacia), digested with Not I and cloned into the Not I  
 and Eco RI sites of the modified pRTD vector. Library  
 went through one round of normalization. Library

ORIGIN constructed by Bento Soares and M. Fatima Bonaldo."

Query Match 33.4%; Score 483.8; DB 6; Length 694;  
 Best Local Similarity 99.4%; Pred. No. 8.2e-116;  
 Matches 496; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

1 ATGAAACCGGGAGTCCCTTTTATGACCTTGTGCTGCTGCACTGCGCTCTCCCA 60  
 196 ATGAAACCGGGAGTCCCTTTTATGACCTTGTGCTGCTGCACTGCGCTCTCCCA 255

61 GCAGCCACTCAGGGAAGAAAGTGTGCTGCGC-AAAAAGGGATACAGTGAATGAC 119  
 256 GCAGCCACTCAGGGAAGAAAGTGTGCTGCGC-AAAAAGGGATACAGTGAATGAC 315

120 CTGACGCTTCCGAGAAAGACATCAATTCACCTGGAATACTCCAAACGATPAA 179  
 316 CTGACGCTTCCGAGAAAGACATCAATTCACCTGGAATACTCCAAACGATPAA 375

180 GATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGC 239  
 376 GATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGC 435

240 TGACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGAACTTTAA 299  
 436 TGACTCAAGAAAGCCTTTGGACCAAGAACTTCCCTGATCATCAAGAACTTTAA 495

300 GATGAGAGACTCAGATCTTCACTGTGTAAGTGAAGACCAAGAAAGAGGTGCATT 359  
 496 GATGAGAGACTCAGATCTTCACTGTGTAAGTGAAGACCAAGAAAGAGGTGCATT 555

360 GCTAGTGTGGATTTGATGCTGCACTGACACCCACTGCTTCAAGGGGAGAGCTGAC 419  
 556 GCTAGTGTGGATTTGATGCTGCACTGACACCCACTGCTTCAAGGGGAGAGCTGAC 615

420 CTGACCTTGGAGAGCCCTTGGTGAAGCCCTCACTGCAATGTAGAGTCCAGGGG 479  
 616 CTGACCTTGGAGAGCCCTTGGTGAAGCCCTCACTGCAATGTAGAGTCCAGGGG 675

480 TAAAAACATACAGGGGGGG 498  
 676 TAAAAACATACAGGGGGGG 694

RESULT 15  
 CO249540 803 bp mRNA linear EST 23-JUN-2004  
 LOCUS  
 DEFINITION AGENCOURT\_26527156 NIH\_MGC\_212 Homo sapiens cDNA clone  
 IMAGE:30921364 5', mRNA sequence.  
 ACCESSION CO249540  
 VERSION CO249540.1 GI:49120610  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 1 (bases 1 to 803)  
 NIH-MGC http://mgi.nci.nih.gov/  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: gcapbs-remail.nih.gov  
 Tissue Procurement: Mary Hendrix  
 cDNA Library Preparation: M. Bento Soares, University of Iowa  
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.lnl.gov  
 Plate: NDAM158 row: 0 column: 05

FEATURES High quality sequence stop: 525.  
 Location/Qualifiers  
 1..803  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:30921364"  
 /tissue\_type="Chondrosarcoma Lung Metastasis cell lines"  
 /lab\_host="DH10B (T1 phage resistant)"  
 /clone\_1ib="NIH MGC 212"  
 /note="Organ: Lung; Vector: pX-Asc; Site: 1: Ecor I;  
 Site: 2: Not I; The library was constructed according  
 Bonaldo, Lennon and Soares, Genome Research, 6:791-806,  
 1996. Denatured RNA was size fractionated on a 1% agarose  
 gel. First strand cDNA synthesis was primed with oligo-dT  
 primer containing a Not I site. Double strand cDNA was  
 size selected according to mRNA size fraction, ligated  
 with Ecor I adaptor, digested with Not I and then cloned  
 directionally into pX-Asc vector. The library tag  
 sequence located between the Not I site and the polyA tail  
 is GATTAAGCCCA. Tissue was provided by Mary Hendrix."

## ORIGIN

Query Match 33.4%; Score 480.6; DB 7; Length 803;  
 Best Local Similarity 96.5%; Pred. No. 5.9e-115;  
 Matches 501; Conservative 0; Mismatches 16; Indels 2; Gaps 1;

1 ATGAAACCGGGAGTCCCTTTTATGACCTTGTGCTGCTGCACTGCGCTCTCCCA 60  
 173 ATGAAACCGGGAGTCCCTTTTATGACCTTGTGCTGCTGCACTGCGCTCTCCCA 232

61 GCAGCCACTCAGGGAAGAAAGTGTGCTGCGC-AAAAAGGGATACAGTGAATGAC 120  
 233 GCAGCCACTCAGGGAAGAAAGTGTGCTGCGC-AAAAAGGGATACAGTGAATGAC 232

121 TGTACAGCTTCCAGAAAGACATCAATTCACCTGGAATACTCCAAACGATPAA 180  
 293 TGTACAGCTTCCAGAAAGACATCAATTCACCTGGAATACTCCAAACGATPAA 352

181 ATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 240  
 353 ATTCTGGGAAATCAGGGCTCTCTTAACTAAAGTCCATCCAGCTGAATGATCGGCT 412

241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAA 300  
 413 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCTGATCATCAAGAACTTTAA 472

301 ATGAGAGCTCAGATCTTCACTGTGTAAGTGAAGACCAAGAAAGAGGTGCATTG 360  
 473 ATGAGAGCTCAGATCTTCACTGTGTAAGTGAAGACCAAGAAAGAGGTGCATTG 532

361 CTAGTGTGGATTTGATGCTGCACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 420  
 533 CTAGTGTGGATTTGATGCTGCACTGACACCCACTGCTTCAAGGGGAGAGCTGACC 592

421 CTGACCTTGGAGAGCCCTTGGTGAAGCCCTCAGTGAATGATGATCCAGGGGT 480  
 593 CTGACCTTGGAGAGCCCTTGGTGAAGCCCTCAGTGAATGATGATGATCCAGGGGT 650

481 AAAAACATACAGGGGGGAGAACCTTCCTCGTGTCTAG 519  
 651 TAAAAACATACAGGGGGGAGAACCTTCCTCGTGTCTAG 689

Search completed: November 22, 2004, 14:13:11  
 Job time : 4905 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 55.6204 Seconds  
(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRFHLVLVQLALLP.....GTRLEIKVPRSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	458	1 RWHUT4	T-cell surface gly
2	916	34.3	432	1 RWCZT4	T-cell surface gly
3	831	31.1	432	1 RWMQT4	T-cell surface gly
4	750	28.1	268	2 A56446	Ig heavy chain V r
5	668	25.0	249	2 S41374	single chain Fv an
6	618.5	22.2	459	2 M46254	CD4 precursor - ra
7	616	22.1	233	2 UC5322	p53 specific singl
8	596	22.3	432	2 S30193	T-cell surface gly
9	508.5	19.1	457	2 A27449	T-cell surface gly
10	507	19.0	128	2 S40343	Ig kappa chain V-J
11	498	18.7	107	2 S34005	Ig kappa chain V r
12	498	18.7	117	2 S40362	Ig kappa chain V r
13	498	18.7	457	1 RWMST4	T-cell surface gly
14	485.5	18.2	116	2 B25555	Ig kappa chain V-I
15	485	18.2	123	2 S40378	Ig kappa chain V-I
16	481.5	18.0	131	2 S40328	Ig kappa chain V-I
17	480.5	18.0	109	1 K3HUPM	Ig kappa chain V-I
18	475.5	17.8	215	2 UB0244	Ig kappa chain V r
19	466	17.5	111	2 S23628	Ig kappa chain V r
20	460	17.2	108	2 G44151	Ig kappa chain V r
21	458	17.2	128	2 A56701	Ig kappa chain V r
22	456.5	17.1	114	2 S54905	Ig kappa chain V r
23	456	17.1	128	2 S40345	Ig kappa chain V-J
24	455	17.1	114	2 S46375	Ig kappa chain V-J
25	455	17.1	129	2 S23627	Ig kappa chain V r
26	454	17.0	115	1 K3HUC8	Ig kappa chain V r
27	453	17.0	98	2 I30608	Ig kappa chain V-I
28	453	17.0	115	2 A30553	Ig kappa chain pre
29	453	17.0	144	2 PL0106	Ig kappa chain pre

30	451	16.9	95	2 PH0868	Ig kappa chain V r
31	451	16.9	215	2 JF0243	Ig kappa chain NIG
32	449	16.8	110	2 S40326	Ig kappa chain V-J
33	448.5	16.8	109	2 H30601	Ig kappa chain V-I
34	448	16.8	128	2 S40379	Ig kappa chain V-J
35	445.5	16.7	627	2 S14683	Ig mu chain precu
36	443.5	16.6	108	2 C30608	Ig kappa chain V-I
37	441.5	16.5	129	2 S46369	Ig light chain V r
38	441.5	16.5	134	2 S38643	Ig kappa chain V r
39	440.5	16.5	107	2 PH0965	Ig kappa chain V r
40	440.5	16.5	109	2 PH0963	Ig kappa chain V r
41	438.5	16.4	109	2 A30608	Ig kappa chain V-I
42	438.5	16.4	109	2 P30607	Ig kappa chain V-I
43	438	16.4	107	2 S57444	Ig kappa chain V-J
44	437.5	16.4	109	2 B30601	Ig kappa chain V-I
45	436.5	16.4	109	2 G30601	Ig kappa chain V-I

## ALIGNMENTS

## RESULT 1

RWHUT4  
T-cell surface glycoprotein CD4 precursor [validated] - human  
N.Alternate names: T-cell surface antigen T4/Leu 3  
C.Species: Homo sapiens (man)  
C.Date: 28-May-1986 #sequence\_revision 31-Dec-1988 #text\_change 09-Jul-2004  
C.Accession: A90872; A32722; A31194; A53287; I54176; I54297; A02109; A30039  
R.Maddon, P.J.; Littman, D.R.; Godfrey, M.; Maddon, D.E.; Chess, L.; Axel, R.  
Cell 42, 93-104, 1985  
A.Title: The isolation and nucleotide sequence of a cDNA encoding the T cell surface pro  
A.Reference number: A90872; MUID:85254948; PMID:2990730  
A.Accession: A90872  
A.Molecule type: mRNA  
A.Residues: 1-25, 'N', 27-458 <MAD>  
A.Cross-references: UNIPROT:P01730  
A.Experimental source: clone pT4B  
R.Littman, D.R.; Maddon, P.J.; Axel, R.  
Cell 55, 541, 1988  
A.Title: Corrected CD4 sequence.  
A.Reference number: A90907; MUID:89028665; PMID:3263213  
A.Contents: annotation; revision to residue 26  
R.Camerini, D.; Seed, B.  
Cell 60, 747-754, 1990  
A.Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
A.Reference number: A32722; MUID:90182664; PMID:2107024  
A.Accession: A32722  
A.Molecule type: mRNA  
A.Residues: 26-426, 428-458 <CAM>  
R.Carr, S.A.; Hemling, M.E.; Folea-Maserman, G.; Sweet, R.W.; Anumula, K.; Barr, J.R.;  
J. Biol. Chem. 264, 21286-21295, 1989  
A.Title: Protein and carbohydrate structural analysis of a recombinant soluble CD4 recep  
A.Reference number: A31934; MUID:90078232; PMID:2592374  
A.Contents: disulfide bonds; carbohydrate-binding sites  
A.Accession: A31934  
A.Molecule type: protein  
A.Residues: 26-394 <CAR>  
R.Liederman, S.; Dewarino, J.A.; Daugherty, B.L.; Foeldvari, I.; Yellin, M.J.; Cleary, A  
Mol. Immunol. 28, 1171-1181, 1991  
A.Title: A single amino acid substitution in a common African allele of the CD4 molecule  
A.Reference number: A53287; MUID:92072595; PMID:1961196  
A.Status: not compared with conceptual translation  
A.Molecule type: mRNA  
A.Residues: 250-264, 'W', 266-280 <LED>  
A.Note: sequence extracted from NCBI backbone (NCBI:68249)  
R.Edwards, M.C.; Gibbs, R.A.  
Genomics 14, 590-597, 1992  
A.Title: A human dimorphism resulting from loss of an Alu.  
A.Reference number: I54176; MUID:93052387; PMID:1330888  
A.Accession: I54176  
A.Status: translated from GR/EMBL/DBJ

A/Molecule type: DNA  
 A/Residues: 1-72 <RB5>  
 A/Cross-references: GB:U47924; GB:M86525; GB:U72506; MID:91633547; PIDN:AA51309.1; PID:  
 R./Hodge, T.W.; Sasso, D.R.; McConugal, J.S.  
 Hum. Immunol. 30, 99-104, 1991  
 A/Title: Humans with OK4-epitope deficiency have a single nucleotide base change in the  
 A/Accession: 154297  
 A/Reference number: 154297; MUID:91216786; PMID:11708753  
 A/Status: translated from GB/EMBL/DBJ  
 A/Molecule type: DNA  
 A/Residues: 1-264, 'W', 266-458 <RB2>  
 A/Cross-references: GB:M35160; MID:9179143; PIDN:AA16069.1; PID:9179144  
 A/Comment: Macrophage tropic strains of HIV-1 bind to a complex of chemokine (C-C) recep  
 C/Genetics:  
 A/Gene: GDB:CD4  
 A/Cross-references: GDB:119767; OMIM:186940  
 A/Map position: 12pter-12p12  
 A/Intons: 16/3  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: AIDS; duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-25/Domain: signal sequence #status predicted <SIG>  
 F/26-458/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
 F/34-111/Domain: immunoglobulin homology <IM1>  
 F/136-186/Domain: immunoglobulin homology #status atypical <IM2>  
 F/216-299/Domain: immunoglobulin homology <IM3>  
 F/321-372/Domain: immunoglobulin homology <IM4>  
 F/397-420/Domain: transmembrane #status predicted <TM>  
 F/421-458/Domain: intracellular #status predicted <INT>  
 F/41-109,155-184,328-370/Disulfide bonds: #status experimental  
 F/286,325/Binding site: carbohydrate (Asn) (covalent) #status experimental

Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 1e-48;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVLTQALLPAATGKRVYLGKGDVLELTCTASQKSIQFHMKNSTQIK 60  
 |||||  
 DB 1 MNRGVPRHLLVLTQALLPAATGKRVYLGKGDVLELTCTASQKSIQFHMKNSTQIK 60  
 |||||  
 QY 61 ILNQGSEFLTKGSPKLNDRSDRSRLMDQGNFPIIKNLKIEBSPYICEVEDQKEEVOL 120  
 |||||  
 DB 61 ILNQGSEFLTKGSPKLNDRSDRSRLMDQGNFPIIKNLKIEBSPYICEVEDQKEEVOL 120  
 |||||  
 QY 121 LVEGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 |||||  
 DB 121 LVEGLTANSDTHLLQGSLTTLTSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
 |||||  
 QY 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
 |||||  
 DB 181 TWTCTVLQNKQKVEFKIDIVLAFQKAS 208  
 |||||

## RESULT 2

RWC274  
 T-cell surface glycoprotein CD4 - chimpanzee  
 N/Alternate names: T-cell surface antigen T4/Leu 3  
 C/Species: Pan troglodytes (chimpanzee)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 09-Jul-2004  
 C/Accession: B32722; A46534  
 R/Camerini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A/Reference number: A32722; MUID:90182664; PMID:2107024  
 A/Accession: B32722  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <CAM>  
 A/Cross-references: UNIPROT:P16004; GB:M31135  
 R./Pomgaard, A.; Hirsch, V.M.; Johnson, P.R.  
 Eur. J. Immunol. 22, 2973-2981, 1992  
 A/Title: Cloning and sequences of primate CD4 molecules: diversity of the cellular recep  
 A/Reference number: A46534; MUID:93046640; PMID:11425921  
 A/Accession: A46534  
 A/Status: not compared with conceptual translation

A/Molecule type: mRNA  
 A/Residues: 3-399 <FOM>  
 A/Note: sequence extracted from NCBI backbone (NCBIP:118332)  
 C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F/1-373/Domain: extracellular #status predicted <EXT>  
 F/9-86/Domain: immunoglobulin homology <IM1>  
 F/111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F/191-274/Domain: immunoglobulin homology <IM3>  
 F/286-347/Domain: immunoglobulin homology <IM4>  
 F/312-395/Domain: transmembrane #status predicted <TM>  
 F/396-432/Domain: intracellular #status predicted <INT>  
 F/16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F/271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 34.3%; Score 916; DB 1; Length 432;  
 Best Local Similarity 97.8%; Pred. No. 4.9e-41;  
 Matches 179; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 26 KRVVLGKGDVLELTCTASQKSIQFHMKNSTQIKILNQGSEFLTKGSPKLNDRSDRS 85  
 |||||  
 DB 1 KRVVLGKGDVLELTCTASQKSIQFHMKNSTQIKILNQGSEFLTKGSPKLNDRSDRS 85  
 |||||  
 QY 86 LMDQGNFPIIKNLKIEBSPYICEVEDQKEEVOLVFGILTANSDTHLLQGSLTTLTLES 145  
 |||||  
 DB 86 LMDQGNFPIIKNLKIEBSPYICEVEDQKEEVOLVFGILTANSDTHLLQGSLTTLTLES 120  
 |||||  
 QY 146 PPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGVTCTVLQNKQKVEFKIDIVLAFQ 205  
 |||||  
 DB 121 PPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSGVTCTVLQNKQKVEFKIDIVLAFQ 180  
 |||||  
 QY 206 KAS 208  
 |||||  
 DB 181 KAS 183  
 |||||

## RESULT 3

RWC274  
 T-cell surface glycoprotein CD4 - rhesus macaque  
 N/Alternate names: T-cell surface antigen T4/Leu 3  
 C/Species: Macaca mulatta (rhesus macaque)  
 C/Date: 30-Sep-1993 #sequence\_revision 30-Sep-1993 #text\_change 16-Jul-1999  
 C/Accession: C32722  
 R/Camerini, D.; Seed, B.  
 Cell 60, 747-754, 1990  
 A/Title: A CD4 domain important for HIV-mediated syncytium formation lies outside the vi  
 A/Reference number: A32722; MUID:90182664; PMID:2107024  
 A/Accession: C32722  
 A/Molecule type: mRNA  
 A/Residues: 1-432 <CAM>  
 A/Cross-references: GB:M31134  
 C/Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells ct  
 C/Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 C/Keywords: duplication; glycoprotein; T-cell; transmembrane protein  
 F/1-432/Product: T-cell surface glycoprotein CD4 #status predicted <MAT>  
 F/1-373/Domain: extracellular #status predicted <EXT>  
 F/9-86/Domain: immunoglobulin homology <IM1>  
 F/111-161/Domain: immunoglobulin homology #status atypical <IM2>  
 F/180-293/Domain: immunoglobulin homology <IM3>  
 F/296-347/Domain: immunoglobulin homology <IM4>  
 F/312-395/Domain: transmembrane #status predicted <TM>  
 F/396-432/Domain: intracellular #status predicted <INT>  
 F/16-84,130-159,303-345/Disulfide bonds: #status predicted  
 F/271,300/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 31.1%; Score 831; DB 1; Length 432;  
 Best Local Similarity 88.0%; Pred. No. 1.3e-36;  
 Matches 161; Conservative 12; Mismatches 10; Indels 0; Gaps 0;

QY 26 KRVVLGKGDVLELTCTASQKSIQFHMKNSTQIKILNQGSEFLTKGSPKLNDRSDRS 85  
 |||||  
 DB 1 KRVVLGKGDVLELTCTASQKSIQFHMKNSTQIKILNQGSEFLTKGSPKLNDRSDRS 85  
 |||||

Db 1 KKVVLGKKGDVTELTCTASQKNTQPFHMKNSNQIKILGILFTKGPSEKLSDRADSRKS 60  
 QY 86 LMDQNFPLIINKLKIEDSDTYICEVEDQKEVQLVREGLTANSPHLLQGSLLTTLLES 145  
 Db 61 LMDQCFSPITIKNLKIEDSDTYICEVENKKEVEILLFGLTANSDTHLLQGSLLTTLLES 120  
 QY 146 PPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSDGTWCTVQONOKVEFKIDIVLAFO 205  
 Db 121 PGGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSDGTWCTVQONOKVEFKIDIVLAFO 180  
 QY 206 KAS 208  
 Db 181 KAS 183

## RESULT 4

A:Accession: A56446  
 Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/c)  
 C:Species: Mus musculus (house mouse)  
 C:Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
 C:Accession: A56446  
 R:Yang, F.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.  
 J. Biol. Chem. 270, 7829-7835, 1995  
 A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical  
 A:Reference number: A56446; MUID:95229583; PMID:7713873  
 A:Accession: A56446  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-268 <TAN>  
 A:Cross-references: GB:U20617  
 C:Keywords: heterotrimer; immunoglobulin

Query Match 28.1%; Score 750; DB 2; Length 268;  
 Best Local Similarity 55.6%; Pred. No. 1.4e-32;  
 Matches 150; Conservative 39; Mismatches 57; Indels 24; Gaps 7;  
 QY 243 SOVQLLESGAEVKKFGSSVKSVCXASG---DTFIRYSTFWROAPGQLEMMGRITTL 298  
 Db 2 AOVKLQESGAEVKKFGASVKSCTTSGRVTKDTYMH---WVKQRPQGLEWIGRIAPAN 57  
 QY 239 DVAAHYAPHLOGRVTTADKSTSTVYLELNLSDDTAVYFCAGVYEGEADGEYDNGFL 358  
 Db 58 GITKIDPKFGKATIAADTSSNTAVLQLSLTSEDTAVYCYASY-----LTRYEN---- 108  
 QY 339 KMGCGTLVTYSSGGSGSGSGSGSGSELELTQSPATLSVSPGERPATLSCTASGSVSD 418  
 Db 109 -YMGCGTLVTYSSGGSGSGSGSGSGSDIELTQSPATLSAGLGEKVTMSCRASSVNF- 166  
 QY 419 LAMVQKKGQAPRLIYGASTRATGVPARFSGSGSGABFTLTSSLSQSEDFAVYCCQYN 478  
 Db 167 IYMWQKSDASPKLWVYTTSHLPFGVPAPFSGSGSGSNSTLSTISSMEGDAATYVCCQPT 226  
 QY 479 NMPRYTGGQTRLEIKLVPRSGGHHHHH 508  
 Db 227 SSP--FTFGSGTKLEIK---RSAHHHHH 251

## RESULT 5

S41374  
 single chain Fv antibody - mouse  
 C:Species: Mus musculus (house mouse)  
 C:Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
 C:Accession: S41374  
 R:Artsaenko, O.; Weiler, E.W.; Muentz, K.; Conrad, U.  
 submitted to the EMBL Data Library, January 1994  
 A:Description: Construction and functional characterization of a single chain Fv antibody  
 A:Reference number: S41374  
 A:Accession: S41374  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-249 <ART>  
 A:Cross-references: EMBL:Z29480

Query Match 25.0%; Score 668; DB 2; Length 249;  
 Best Local Similarity 52.9%; Pred. No. 2.5e-28;  
 Matches 136; Conservative 40; Mismatches 65; Indels 16; Gaps 4;  
 QY 244 QVQLLESGAEVKKFGSSVKSVCXASGDTFIRYSTFWROAPGQLEMMGRITTLIDVAY 303  
 Db 1 QVQLQSGSELVPRGASVYKLSCTAGSNFKODYIMHWQRPKEGLWIAIAPASGVNYK 60  
 QY 304 APHLQKRVTTADKSTSTVYLELNLSDDTAVYFCAGVYEGEADGEYDNGFLKMGQ 363  
 Db 61 VPRFQDKATITADTSSNTAVLQLSLTSEDTAVYCYA-----RDITLYTSIGY--WGQ 111  
 QY 364 GTLVTVTSGSGSGSGSGSGSELELTQSPATLSVSPGERPATLSCTASGSV-----SSD 418  
 Db 112 GSTVTVSSRGGSGSGSGSGSDIELTQSPSVVIRGESVYSISCRSSKLLSYSDGSY 171  
 QY 419 LAMVQKKGQAPRLIYGASTRATGVPARFSGSGSGABFTLTSSLSQSEDFAVYCCQYN 478  
 Db 172 LFMFLQRBQSGPQLLIYMSNLASGVPRFSGSGSGFTTIRISRVEADGVYVCMQHR 231  
 QY 479 NMPRYTGGQTRLEIK 495  
 Db 232 EYP--LTFGAGTKLEIK 246

## RESULT 6

A:Accession: A46254  
 CD4 precursor - rabbit  
 C:Species: Oryctolagus cuniculus (domestic rabbit)  
 C:Date: 21-Sep-1993 #sequence\_revision 18-Nov-1994 #text\_change 09-Jul-2004  
 C:Accession: A46254  
 R:Hague, B.F.; Sawadikoseol, S.; Brown, T.U.; Lee, K.; Recker, D.P.; Kindt, T.J.  
 Proc. Natl. Acad. Sci. U.S.A. 89, 7963-7967, 1992  
 A:Title: CD4 and its role in infection of rabbit cell lines by human immunodeficiency vi  
 A:Note: sequence extracted from NCBI backbone (NCBI:112732, NCBI:112733)  
 C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
 F:322-372/Domain: immunoglobulin homology <IMM>

Query Match 23.2%; Score 618.5; DB 2; Length 459;  
 Best Local Similarity 59.6%; Pred. No. 1.7e-25;  
 Matches 127; Conservative 34; Mismatches 47; Indels 5; Gaps 2;  
 QY 1 MNRGVPFRLILVLDLALLPAATGKRVLGKGDVTELTCTASQKSIQFHMKNSNQIK 60  
 Db 1 MNRRIYFQCLLLVPLALLPAATGKTVVRGAGIYELPCOSSQKRSVFNWKKANQVK 60  
 QY 61 ILGNNG-----SFLTKGPSEKLNDRADSRSLMDQGNFPLIINKLKIEDSDTYICEVEDQKE 116  
 Db 61 ILGNQSSSSSFWLKGNSPLNRYBSKKMMQDSFPLVITKDLKMDSDGTTICEVGDKKM 120  
 QY 117 EVQLLVFGLTANSDTHLLQGSLLTTLLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLEL 176  
 Db 121 EVELLVFRLTANPMTRLHLHGQSLTLLEGPVSVPQWPKPKKTIETGTCMPKRL 180  
 QY 177 ODGSGWTCTV-LQNOKVEFKIDIVLAFOKAS 208  
 Db 181 QDSGWSCHLSFODQNKLEIDIKITIVGFPRAS 213

## RESULT 7

JC5322  
 p53 specific single-chain antibody Pab421 - human  
 C:Species: Homo sapiens (man)  
 C:Date: 15-May-1997 #sequence\_revision 15-May-1997 #text\_change 18-Jul-1997  
 C:Accession: JC5322  
 R:Janot, C.B.; Hynes, N.E.  
 Biochem. Biophys. Res. Commun. 230, 242-246, 1997



Db 81 RFGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP--YTFGGGTKEIK 127

## RESULT 11

S34005  
Ig kappa chain V region - human  
C:Species: Homo sapiens (man)  
C>Date: 02-Dec-1993 #sequence\_revision 10-Nov-1995 #text\_change 21-Jan-2000  
C:Accession: S34005; S30524  
R:Marlette, X.; Tsapis, A.; Brouet, J.C.  
Eur. J. Immunol. 23, 846-851, 1993  
A>Title: Nucleotide sequence analysis of the variable domains of four human monoclonal  
A:Reference number: S34001; MUID:93209281; PMID:7681398  
A:Accession: S34005  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-107 <MAR>  
A:Cross-references: EMBL:Z18330  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
P:16-90/Domain: immunoglobulin homology <IMM>

Query Match 18.7%; Score 498; DB 2; Length 107;  
Best Local Similarity 88.1%; Pred. No. 8.2e-20;  
Matches 96; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

QY 387 ELETQSPATLSVSPERATLSCRASVSVDLAWQKPGQAPRLIYGASTRATGVA 446

Db 1 EIVMTQSPATLSVSPERATLSCRASVSVDLAWQKPGQAPRLIYGASTRATGIPA 60

QY 447 RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWPRTYFGGGTKEIK 495  
Db 61 RFGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP--RTFGGGTKEIK 107

## RESULT 12

S40362  
Ig kappa chain - human  
C:Species: Homo sapiens (man)  
C>Date: 06-Mar-1994 #sequence\_revision 26-May-1995 #text\_change 21-Jan-2000  
C:Accession: S40362  
R:Klein, R.; Jaenichen, R.; Zachau, H.G.  
Eur. J. Immunol. 23, 3248-3271, 1993  
A>Title: Expressed human immunoglobulin chi genes and their hypermutation.  
A:Reference number: S40312; MUID:94080891; PMID:8258341  
A:Accession: S40362  
A:Status: preliminary; translation not shown  
A:Molecule type: mRNA  
A:Residues: 1-117 <KLB>  
A:Cross-references: EMBL:X72472; NID:g441412; PID:g441413  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotrimer; immunoglobulin  
P:26-100/Domain: immunoglobulin homology <IMM>

Query Match 18.7%; Score 498; DB 2; Length 117;  
Best Local Similarity 88.1%; Pred. No. 8.9e-20;  
Matches 96; Conservative 7; Mismatches 4; Indels 2; Gaps 1;

QY 387 ELETQSPATLSVSPERATLSCRASVSVDLAWQKPGQAPRLIYGASTRATGVA 446

Db 11 EIVMTQSPATLSVSPERATLSCRASVSVDLAWQKPGQAPRLIYGASTRATGIPA 70

QY 447 RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNNWPRTYFGGGTKEIK 495  
Db 71 RFGSGSGTEFTLTISLQSEDPFAYVYCCQYNNWP--LTFGGGTKEIK 117

## RESULT 13

RHM5T4  
T-cell surface glycoprotein CD4 precursor - mouse  
N:Alternate names: T-cell differentiation antigen L3T4; T-cell surface antigen T4/Leu 3  
C:Species: Mus musculus (house mouse)

C>Date: 30-Jun-1987 #sequence\_revision 30-Jun-1987 #text\_change 09-Jul-2004  
C:Accession: A02110; A26038; A39893; A39955; I54564; I69018; A47642  
R:Tourville, B.; Gorman, S.D.; Field, E.H.; Hunziker, T.; Parner, J.R.  
Science 234, 610-614, 1986  
A>Title: Isolation and sequence of L3T4 complementary DNA clones: expression in T cells  
A:Reference number: A02110; MUID:87018845; PMID:3094146  
A:Accession: A02110  
A:Molecule type: mRNA

A:Residues: 1-457 <TOU>  
A:Cross-references: UNIPROT:P06332; GB:M13816; NID:G192070; PIDN:AAA37267.1; PID:G309112  
R:Riltman, D.R.; Gettner, S.N.  
Nature 325, 453-455, 1987  
A>Title: Unusual insertion in the immunoglobulin domain of the newly isolated murine CD4 (L  
A:Reference number: A26038; MUID:87115821; PMID:3027575  
A:Accession: A26038  
A:Molecule type: mRNA

A:Residues: 1-457 <LIT>  
A:Cross-references: GB:X04836; NID:G50353; PIDN:CAA28539.1; PID:G50354  
R:Gorman, S.D.; Tourville, B.; Parner, J.R.  
Proc. Natl. Acad. Sci. U.S.A. 84, 7644-7648, 1987

A>Title: Structure of the mouse gene encoding CD4 and an unusual transcript in brain.  
A:Reference number: A39893; MUID:88041159; PMID:2823269  
A:Accession: A39893  
A:Molecule type: DNA

A:Residues: 1-25, 'E', '27-457 <GOR>  
A:Cross-references: GB:M17080; GB:J03003; NID:G192515; PIDN:AAA37402.1; PID:G387124  
R:Maddon, P.J.; Moineaux, S.M.; Maddon, D.E.; Zimmerman, K.A.; Godfrey, M.; Alt, F.W.;  
Proc. Natl. Acad. Sci. U.S.A. 84, 9155-9159, 1987

A>Title: Structure and expression of the human and mouse T4 genes.  
A:Reference number: A39955; MUID:88097446; PMID:3501122  
A:Accession: A39955

A:Status: nucleic acid sequence not shown; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 25-457 <MAD>  
A>Note: the cited Genbank accession number, J03564, is not in release 101.0

R:Parner, J.R.; Hunziker, T.  
Immunol. Rev. 100, 109-127, 1987  
A>Title: L3T4 and the immunoglobulin gene superfamily: New relationships between the imm

A:Reference number: I54564; MUID:88152875; PMID:3326818  
A:Accession: I54564  
A:Status: translated from GB/EMBL/DBJ

A:Molecule type: mRNA  
A:Residues: 1-457 <RES>  
A:Cross-references: GB:M6850; NID:G198670; PIDN:AAA39401.1; PID:G198671

A:Accession: I69018  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: DNA

A:Residues: 208-318 <RE2>  
A:Cross-references: GB:M6851; NID:G198672; PIDN:AAA39402.1; PID:G554183  
R:Classon, B.J.; Tsagaratos, J.; Kirsbaum, L.; Maddox, J.; Mackay, C.R.; Brandon, M.; M

Immunogenetics 23, 129-132, 1986  
A>Title: The L3T4 antigen in mouse and the sheep equivalent are immunoglobulin-like.  
A:Reference number: A47642; MUID:86166694; PMID:3082751  
A:Accession: A47642

A:Molecule type: protein  
A:Residues: 27-43 <CLA>  
C:Comment: This protein is expressed on most thymocytes, on a subset of mature T-cells t

C:Genetics:  
A:Introns: 18/1; 74/1; 128/1; 207/1; 319/1; 386/1; 425/3; 448/2  
C:Superfamily: T-cell surface glycoprotein CD4; immunoglobulin homology  
C:Keywords: alternative initiators; duplication; glycoprotein; T-cell; transmembrane pro

F:11-26/Domain: signal sequence #status predicted <SIG>  
F:127-457/Product: T-cell surface glycoprotein CD4 #status experimental <MAT>  
F:135-114/Domain: immunoglobulin homology <IMM>  
F:139-190/Domain: immunoglobulin homology #status atypical <IM2>

F:220-301/Domain: immunoglobulin homology #status atypical <IM3>  
F:241-457/Product: CD4, brain-specific short form #status predicted <BRA>  
F:321-372/Domain: immunoglobulin homology <IM4>

F:395-419/Domain: transmembrane #status predicted <TMM>  
F:420-457/Domain: intracellular #status predicted <INT>  
F:442-112,159-188,328-370/Disulfide bonds: #status predicted

P:187,298,333,392/Binding site: carbohydrate (Aan) (covalent) #status predicted

Query Match	18.7%	Score 498	DB 19	Length 457
Best Local Similarity	51.6%	Pred. No. 3.3e-19		
Matches	110	Conservative	35	Mismatches 62, Indels 6, Gaps 5
Qy	1	MNRGVPFPH-LTLVLTALLPAAITGGKRVV-LGKKGDITVETCTASQKTSIOFHMKNSNOI		59
Db	1	MCRAISLRRLULLLTQLTQSOLAIVTGGKTLVIGKEGESABLPCESSOKKITVPTWYFSDQR		60
Qy	60	KILNQG--SPLTKG--PSKLNDRASRSLSMGGGFPILIKLTKEDSDTVCEVDDKE		116
Db	61	KILGGHGCVLIRGSGPSQP--DRPSKKGAMWKGSPPLINKLKREDSDTYICELENRKE		119
Qy	117	EVQLLVFGLTANSDTHLQGGSLTLTLBS--PPGSSPVQCRSPRKNIQGGKTLTSVSOLE		175
Db	120	EVELVWFVKTSPGPTSLQGGSLTLTLTSSNSKVSNPRLFECHKRKGKRVVSGSKVLSMSNLR		179
Qy	176	LQDSGTWTCYLVQNGKVKVEFKIDIVYLAFOKRS		208
Db	180	VQDSDFMNCITVLTLDQKKWFGMTSLVVGQSTRA		212

```

RESULT 14
B26555
Ig kappa chain V-III region (Ger) - human
C:Species: Homo sapiens (man)
C:Date: 05-Jun-1988 #sequence_revision 05-Jun-1988 #text_change 21-Jan-2000
C:Accession: B26555
R:Madaugh, C.R./ Litman, G.W.
J. Biol. Chem. 262, 3671-3673, 1987
A>Title: Atypical glycosylation of an IgG monoclonal cryoimmunoglobulin.
A:Reference number: A92830, PMID:87137666, PMID:3102493
A:Accession: B26555
A:Molecule type: protein
A:Residues: 1-116 <MID>
C:Superfamily: immunoglobulin V region; immunoglobulin homology
C:Keywords: heterotrimer; immunoglobulin
F:1-90/Domain: immunoglobulin homology <IMM>

```

	Query Match	Similarity	85.2%	Score	485.5	DB 2	Length	116
Best local			85.3%	Pred. No.	46-19			
Matches	93	Conservative	9	Mismatches	6	Indels	1	Gaps
QY	387	EELTQSPATLTVSNGERATLSCRASESVISDLANYQQKPGQAPRLIIYGASTRATQVPA					446	
DB	1	DIWMTQVPTLTVSNGERATLSCRAEQSVISDLANYQQKPGQAPRLIIYGASTRATQVPA					60	
QY	447	RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNNMPRYTEGQTRLEIK					495	
DB	61	RFGSGSGAEFTLTISLQSEDPFAYVYCCQYNDWPP-IRFGGTRLEIK					108	

```

RESULT 15
S40378
Ig kappa chain - human
C|Species: Homo sapiens (man)
C|Date: 06-Mar-1994 #sequence_revision 26-May-1995 #text_change 21-Jan-2000
C|Accession: S40378
R|Klein, R.; Jaenichen, R.; Zachau, H.G.
Eur. J. Immunol. 23, 3246-3271, 1993
A|Title: Expressed human immunoglobulin chi genes and their hypermutation.
A|Reference number: S40312, MUID:94080891, PMID:8258341
A|Accession: S40378
A|Status: preliminary; translation not shown
A|Molecule type: mRNA
A|Residues: 1-123 <KLE>
A|Cross-references: EMBL:X72488, NID:g441444, PID:CAA51156.1, PID:g441445
A|Superfamily: Immunoglobulin V region, immunoglobulin homology
C|Keywords: heterotetramer, immunoglobulin
C|Keywords: Domain: immunoglobulin homology <IMM>

```

Query Match	18.2%	Score 485;	DB 2;	Length 123;
Best Local Similarity	80.4%	Pred. No. 4.4e-19;		
Matches	90;	Conservative	10;	Mismatches 12;
				Indels 0;
				Gaps 0;

QY 381 GSELELTOSPALSYSPERATLSCRASESYSGLAYOOKPCQAPRLILYGSTATG 443  
 Db 11 GTGEIWTGSPALSYSPEDRTLSCRASSGYSILAYQYRPGQAPRLILYGSTATAG 70  
 QY 444 VPARFGSGGSAFFLTITISLQSDPAIYYCOQYNNMPPRTFGQGRLEIK 495  
 Db 71 IPARFSSGSGEFTLTITISLQSDAPAIYYCOQYIDMPPMTFGGRILDIK 122

Search completed: November 21, 2004, 13:43:29  
Job time : 57.6204 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:28:31 ; Search time 281.81 Seconds  
(without alignments)  
1037.189 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRHLLVLQLALLP.....GTRLEIKLVPRGSHHHHH 508

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues

1825181

Total number of hits satisfying chosen parameters:

1825181

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : UniProt 02:\*

1: uniprot\_sprot:\*  
2: uniprot\_crembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1064	39.9	458	1	CD4_HUMAN
2	1040	39.0	458	1	CD4_PANTR
3	953	35.7	458	1	CD4_MACFA
4	951	35.6	458	1	CD4_MACFU
5	945	35.4	458	1	CD4_MACMU
6	944	35.4	458	1	CD4_MACNE
7	926	34.7	458	1	CD4_CERAB
8	820	30.7	397	1	CD4_CERTO
9	819	30.7	397	1	CD4_ERIPA
10	803	30.1	397	2	009261
11	803	30.1	397	2	009262
12	803	30.1	397	2	009263
13	803	30.1	397	2	009264
14	799	29.9	397	2	009259
15	799	29.9	397	2	009260
16	778.5	29.1	457	2	08H2T8
17	777.5	29.1	457	2	08H2T7
18	774.5	29.0	457	1	CD4_SAIISC
19	766	28.7	255	2	06K605
20	766	28.7	255	2	CAG34081
21	750	28.1	243	2	Q7T0M2
22	734	27.5	298	2	Q9GYF0
23	685.5	25.7	241	2	0921A6
24	632	23.7	463	1	CD4_CAMFA
25	630	23.6	455	2	0710E2
26	630	23.6	455	2	AAQ03208
27	625	23.4	455	2	Q9X578
28	618.5	23.2	459	1	CD4_RABIT
29	612	22.3	218	2	Q925S1
30	596	22.3	432	2	06L8K1
31	596	22.3	432	2	CAB37664

32	590.5	22.1	474	2	P79355	P79355 felis silve
33	585	21.9	417	2	AA99809	AA99809 sus scrofa
34	585	21.9	457	2	Q6GYR3	Q6GYR3 sus scrofa
35	585	21.9	457	2	AA99808	AA99808 sus scrofa
36	563	21.1	406	2	Q6R3N3	Q6R3N3 sus scrofa
37	563	21.1	447	2	Q6R3N4	Q6R3N4 sus scrofa
38	508.5	19.1	457	1	CD4_RAT	P05540 rattus norv
39	505.5	18.9	129	1	KV3H_HUMAN	P04207 homo sapien
40	498	18.7	457	1	CD4_MOUSE	P06332 mus musculu
41	490	18.4	457	2	Q61396	Q61396 mus musculu
42	485.5	18.2	109	2	Q9UL85	Q9UL85 homo sapien
43	480.5	18.0	109	1	KV3F_HUMAN	P01624 homo sapien
44	480	18.0	108	2	Q9UL83	Q9UL83 homo sapien
45	480	18.0	234	2	AAH30813	AAH30813 homo sapi

## ALIGNMENTS

RESULT 1	CD4_HUMAN	STANDARD	PRT	458 AA.
AC	P01730;			
DT	21-JUL-1986 (Rel. 01, Created)			
DT	01-NOV-1988 (Rel. 09, Last sequence update)			
DT	01-OCT-2004 (Rel. 45, Last annotation update)			
DE	T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).			
GN	Name=CD4;			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OK	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=85254948; PubMed=2990730;			
RT	Maddon P.J., Littman D.R., Godfrey M., Maddon D.B., Chess L., Axel R.,			
RT	"The isolation and nucleotide sequence of a cDNA encoding the T cell			
RT	surface protein T4: a new member of the immunoglobulin gene family.";			
RL	Cell 42:93-104 (1985).			
RN	[2]			
RP	REVISION TO 26.			
RA	MEDLINE=89028665; PubMed=3263213;			
RT	Littman D.R., Maddon P.J., Axel R.;			
RT	"Corrected CD4 sequence.";			
RL	Cell 55:541-541 (1988).			
RN	[3]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=96303695; PubMed=8723724;			
RT	Anselmi-Lari M.A., Muzny D.M., Lu J., Lu F., Lilley C.E., Spanos S.,			
RT	Walley T., Gibbs R.A.;			
RT	"A gene-rich cluster between the CD4 and triosephosphate isomerase			
RT	genes at human chromosome 12p13.";			
RL	Genome Res. 6:314-326 (1996).			
RN	[4]			
RP	SEQUENCE FROM N.A., AND VARIANT TRP-265.			
RA	MEDLINE=91216786; PubMed=1708753;			
RT	Hodge T.W., Sasse D.R., McDougal J.S.;			
RT	"Humans with OKT4-epitope deficiency have a single nucleotide base			
RT	change in the CD4 gene, resulting in substitution of TRP240 for			
RT	ARG240.";			
RL	Hum. Immunol. 30:99-104 (1991).			
RN	[5]			
RP	SEQUENCE FROM N.A.			
RA	TISUS-Panceraas;			
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;			
RT	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RT	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,			
RT	Altschuler S.F., Zeeberg B., Buelow K.H., Schaefer C.F., Bhat N.K.,			
RT	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,			
RT	Diatchenko L., Marutina K., Farmer A.A., Rubin G.M., Hong L.,			
RT	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RT	Brownstein M.J., Ueidi T.B., Toshlyuk S., Carninci P., Prange C.,			

RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.S., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Wozley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
 RA Vallatton D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Whiting J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,  
 RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.,  
 RT "generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RN [6]  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921.  
 RA Pomegaard A., Hirsch V.M., Johnson P.R.,  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [7]  
 RP SEQUENCE OF 26-394.  
 RX MEDLINE=90078232; PubMed=2592374;  
 RA Carr S.A., Hemling M.E., Folena-Wasserman G., Sweet R.W., Annumula K.,  
 RA Barr J.R., Huddleston M.J., Taylor P.,  
 RT "Protein and carbohydrate structural analysis of a recombinant soluble  
 RT CD4 receptor by mass spectrometry.";  
 RL J. Biol. Chem. 264:21286-21295(1989).  
 RN [8]  
 RP SEQUENCE OF 26-40.  
 RA Zhang Z., Henzel W.,  
 RT "Signal peptide prediction based on analysis of experimentally  
 RT verified cleavage sites.";  
 RL Submitted (JUN-2004) to Swiss-Prot.  
 RN [9]  
 RP X-RAY CRYSTALLOGRAPHY (2.4 ANGSTROMS) OF 26-208.  
 RX MEDLINE=91061881; PubMed=1701030;  
 RA Wang J., Yan Y., Garrett T.P., Liu J., Rodgers D.W., Garlick R.L.,  
 RA Tarr G.E., Hsuan Y., Reinherz E.L., Harrison S.C.,  
 RT "Atomic structure of a fragment of human CD4 containing two  
 RT immunoglobulin-like domains.";  
 RL Nature 348:411-418(1990).  
 RN [10]  
 RP X-RAY CRYSTALLOGRAPHY (2.3 ANGSTROMS) OF 26-208.  
 RX MEDLINE=91061882; PubMed=2247146;  
 RA Ryu S.-E., Kwong P.D., Truneh A., Porter T.G., Arthos J.,  
 RA Rosenberg M., Dai X., Xuong N.-H., Axel R., Sweet R.W.,  
 RA Hendrickson W.A.,  
 RT "Crystal structure of an HIV-binding recombinant fragment of human  
 RT CD4.";  
 RL Nature 348:419-426(1990).  
 RN [11]  
 RP X-RAY CRYSTALLOGRAPHY (3.9 ANGSTROMS) OF 26-388.  
 RX MEDLINE=97311402; PubMed=9168119;  
 RA Wu H., Kwong P.D., Hendrickson W.A.,  
 RT "Dimeric association and segmental variability in the structure of  
 RT human CD4.";  
 RL Nature 387:527-530(1997).  
 RN [12]  
 RP PALMITOYLATION.  
 RX MEDLINE=92317088; PubMed=1618861;  
 RA Crise B., Rose J.K.,  
 RT "Identification of palmitoylation sites on CD4, the human  
 RT immunodeficiency virus receptor.";  
 RL J. Biol. Chem. 267:13593-13597(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.

CC -1- DATABASE: NAME=PROV; NOTE=CD guide CD4 entry;  
 CC WWW=http://www.ncbi.nlm.nih.gov/prov/cd/cd4.htm".  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.iesb-sib.ch/announce/  
 CC or send an email to license@sib-sib.ch).  
 CC -----  
 DR EMBL: M12807; AAA35572.1; -.  
 DR EMBL: U47924; AAB51309.1; -.  
 DR EMBL: M35160; AAA16069.1; -.  
 DR EMBL: BC025782; AAR25782.1; -.  
 DR PIR: A90872; RWHUT4.  
 DR PDB: 1CDH; X-ray; @=26-203.  
 DR PDB: 1CDI; X-ray; @=25-203.  
 DR PDB: 1CDJ; X-ray; @=26-203.  
 DR PDB: 1CDU; X-ray; @=26-203.  
 DR PDB: 1CDY; X-ray; @=26-203.  
 DR PDB: 1G9M; X-ray; C=26-210.  
 DR PDB: 1G9N; X-ray; C=26-210.  
 DR PDB: 1GCI; X-ray; C=26-210.  
 DR PDB: 1JL4; X-ray; D=26-203.  
 DR PDB: 1Q68; NMR; A=421-458.  
 DR PDB: 1WBR; NMR; @=427-445.  
 DR PDB: 1WIP; X-ray; A/B=26-388.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 1WIO; X-ray; A/B=26-388.  
 DR PDB: 3CD4; X-ray; @=26-207.  
 DR IntAct: P01730; -.  
 DR GlycoSuiteDB: P01730; -.  
 DR Genew: HGNC:1678; CD4.  
 DR MIM: 186940; -.  
 DR GO: GO:0005886; C:Plasma membrane; TAS.  
 DR GO: GO:0042101; C:T-cell receptor complex; NAS.  
 DR GO: GO:0015026; F:coreceptor activity; NAS.  
 DR GO: GO:0015029; F:internalization receptor activity; TAS.  
 DR GO: GO:0042289; F:MHC class II protein binding; NAS.  
 DR GO: GO:0004888; F:transmembrane receptor activity; TAS.  
 DR GO: GO:0006955; P:immune response; NAS.  
 DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; NAS.  
 DR GO: GO:0030217; P:T-cell differentiation; NAS.  
 DR GO: GO:0045058; P:T-cell selection; NAS.  
 DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; NAS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR000973; CD4\_TCRg.  
 DR InterPro: IPR007110; Ig\_1like.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_2.  
 DR PRINTS: PR00692; CD4TCANTIGEN.  
 DR SMART: SM00406; IGV\_1.  
 DR PROSITE: PS50835; IG\_1like; 1.  
 KW 3D-structure; Direct protein sequencing; Glycoprotein;  
 KW Immune response; Immunoglobulin domain; Lipoprotein; Palmitate;  
 KW Polymorphism; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 296  
 FT CARBOHYD 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 By similarity.

FT LipID 419 419 S-palmitoyl cysteine.  
 FT LipID 422 422 S-palmitoyl cysteine.  
 Query Match 39.9%; Score 1064; DB 1; Length 458;  
 Best Local Similarity 100.0%; Pred. No. 6.2e-59;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOIK 60  
 DB 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPFLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPFLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGILTANSDTHLQGGSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGILTANSDTHLQGGSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONQKVEFKIDIVVLAFOKAS 208  
 DB 181 TWTCTVLONQKVEFKIDIVVLAFOKAS 208

RESULT 2  
 CD4\_PANTR STANDARD; PRT; 458 AA.  
 AC P16004;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).  
 GN Name=CD4;  
 OS Pan troglodytes (Chimpanzee).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Pan.  
 OX NCBI\_TaxID=9598;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90182664; PubMed=2107024;  
 RA Cameroni D., Seed B.;  
 RT "A CD4 domain important for HIV-mediated syncytium formation lies outside the virus binding site.";  
 RL Cell 60:747-754 (1990).  
 RN [2]  
 RP SEQUENCE OF 26-424 FROM N.A.  
 RC TISSUE=Blood;  
 RX MEDLINE=93049640; PubMed=1425921;  
 RA Pomegaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the cellular receptor for simian immunodeficiency virus/human immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981 (1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL, M31135; AA035407.1; -  
 CC EMBL, X73353; CA051749.1; -  
 CC PIR, B32722; RWC274.  
 CC HSP, P01730; IWI0.

DR GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO:0015025; P:coreceptor activity; ISS.  
 DR GO:0042289; P:MHC class II protein binding; ISS.  
 DR GO:0006955; P:immune response; ISS.  
 DR GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO:0030217; P:T-cell differentiation; ISS.  
 DR GO:0045058; P:T-cell selection; ISS.  
 DR GO:0007165; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro: IPR008424; CD2.  
 DR InterPro: IPR00973; CD4\_TcAg.  
 DR InterPro: IPR007110; Ig\_Like.  
 DR InterPro: IPR003596; Ig\_V.  
 DR Pfam: PF00047; Ig\_2.  
 DR SMART: SM00406; IGV\_1.  
 DR PROSITE: PS00835; IG\_LIKE; 1.  
 DR KX Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25 By similarity.  
 FT CHAIN 26 458 T-cell surface glycoprotein CD4.  
 FT DOMAIN 26 396 Extracellular (Potential).  
 FT TRANSMEM 397 418 Cytoplasmic (Potential).  
 FT DOMAIN 419 458 Ig-like V-type.  
 FT DOMAIN 126 203 Ig-like C2-type 1.  
 FT DOMAIN 204 317 Ig-like C2-type 2.  
 FT DOMAIN 318 374 Ig-like C2-type 3.  
 FT CARBOHYD 296 296 N-linked (GlcNAc. . .) (By similarity).  
 FT CARBOHYD 325 325 N-linked (GlcNAc. . .) (By similarity).  
 FT DISULFID 41 109 By similarity.  
 FT DISULFID 155 184 By similarity.  
 FT DISULFID 328 370 By similarity.  
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).  
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).  
 FT CONFLICT 62 62 L -> S (in Ref. 2).  
 FT CONFLICT 191 191 K -> N (in Ref. 2).  
 SQ SEQUENCE 458 AA; 51057 MW; A7C3AC8A5257D3AD CRC64;

Query Match 39.9%; Score 1040; DB 1; Length 458;  
 Best Local Similarity 98.1%; Pred. No. 2e-57;  
 Matches 204; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOIK 60  
 DB 1 MNRGVPFRLHLVLTALLPATQGGKVVGLGKKGDVTELCTASQKKSIOFHWKNSNOIK 60  
 QY 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPFLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILGNQGSFLTKGSPSKLNRADSRSLMDQGNPFLIKKLIKEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGILTANSDTHLQGGSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGILTANSDTHLQGGSLTTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONQKVEFKIDIVVLAFOKAS 208  
 DB 181 TWTCTVLONQKVEFKIDIVVLAFOKAS 208

RESULT 3  
 CD4\_MACFA STANDARD; PRT; 458 AA.  
 ID CD4\_MACFA  
 AC P79185;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen T4/Leu-3).  
 GN Name=CD4;  
 OS Macaca fascicularis (Crab eating macaque) (Cynomolgus monkey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;

CC Cercopithecinae; Macaca.  
 ON NCBI\_Taxid=9543;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Thymocytes;  
 RA Tatusumi M., Yabe M., Yamada Y.K.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like C2-type domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63349; BAA09673.1; -.  
 DR HSSP; P01730; IWR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 FT SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;  
 SQ  
 Query March 35.7%; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.9%; Pred. No. 6e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

QY 121 LVFGLTANSDPTLLHLOQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVSLQLQDSG 180  
 DB 121 LVFGLTANSDPTLLHLOQSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLTSSVSLQLQDSG 180  
 QY 181 TWTCVTYLOQKQKVEPKIDIVLAFOKAS 208  
 DB 181 TWTCVTYLOQKQKVEPKIDIVLAFOKAS 208  
 RESULT 4  
 CD4\_MACFU STANDARD; PRT; 458 AA.  
 ID CD4\_MACFU  
 AC P79184;  
 DT 15-JUN-1998 (Rel. 36, Created)  
 DT 15-JUN-1998 (Rel. 36, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen  
 T4/Leu-3).  
 GN Name=CD4;  
 OS Macaca fuscata fuscata (Japanese macaque).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopithecinae; Macaca.  
 ON NCBI\_Taxid=9543;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Hashimoto O., Tatusumi M.;  
 RL Submitted (FEB-1997) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (by similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; D63348; BAA09672.1; -.  
 DR HSSP; P01730; IWR.  
 DR GO; GO:0042101; C:T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4\_TcAg.  
 DR InterPro; IPR007110; Ig-Like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 2.  
 DR PRINTS; PR00692; CD4TCANTIGEN.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 126 125  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CARBOHYD 42 42  
 FT CARBOHYD 296 296  
 FT CARBOHYD 325 325  
 FT DISULFID 41 109  
 FT DISULFID 155 184  
 FT DISULFID 328 370  
 FT LIPID 419 419  
 FT LIPID 422 422  
 FT SEQUENCE 458 AA; 50872 MW; 9105479FB5C56FF7 CRC64;  
 SQ  
 Query March 35.7%; Score 953; DB 1; Length 458;  
 Best Local Similarity 88.9%; Pred. No. 6e-52;  
 Matches 185; Conservative 12; Mismatches 11; Indels 0; Gaps 0;

```

FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 335 335 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT DISULFID 370 370 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
SQ SEQUENCE 458 AA; 50828 MW; 76B3E7F08185535 CRC64;

Query Match 35.6%; Score 951; DB 1; Length 458;
Best Local Similarity 88.5%; Pred. No. 8e-52;
Matches 184; Conservative 13; Mismatches 11; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLVLTQALLPATQKGVVLRKSDPTVELTCTASQKSIQPHMKNSNQIK 60
DB 1 NMRGVPFRHLVLTQALLPATQKGVVLRKSDPTVELTCTASQKSIQPHMKNSNQIK 60
QY 61 ILGNQSPFLTKGPSKLNDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDEKQEVLT 120
DB 61 ILGIGSPFLTKGPSKLNDRADSRSLMDQGNFPLIKNIKIEDSDTYICEVEDEKQEVLT 120
QY 121 LVFGITANSDFHLQGSITLTLESPPGSSPSVQCRSPGKNIQGGKTVLSQLELDQSG 180
DB 121 LVFGITANSDFHLQGSITLTLESPPGSSPSVQCRSPGKNIQGGKTVLSQLELDQSG 180
QY 181 TWTCTVLONOKVPEFKIDIVLAFQKAS 208
DB 181 TWTCTVLONOKVPEFKIDIVLAFQKAS 208
QY 181 TWTCTVLONOKVPEFKIDIVLAFQKAS 208
DB 181 TWTCTVLONOKVPEFKIDIVLAFQKAS 208

RESULT 5
CD4_MACMU STANDARD; PRT; 458. AA.
ID CD4_MACMU STANDARD; PRT; 458. AA.
AC P16003; Q29617;
DT 01-APR-1990 (Rel. 14, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca mulatta (rhesus macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Macaca.
OX NCBI_TaxId=9544;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90182664; PubMed=2107024;
RA Camerini D.; Seed B.;
RT "A CD4 domain important for HIV-mediated syncytium formation lies
RT outside the virus binding site.";
RL Cell 60:747-754(1990).
RN [2]
RP SEQUENCE FROM N.A.
RX TISSUE=Thymocytes;
RA Hashimoto O.; Tatsumi M.;
RT "Molecular cloning and expression of macaque CD4s.";
RL Submitted (JUN-1995) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 28-424 FROM N.A.
RX TISSUE=Blood;
RC MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A.; Hirsch V.M.; Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus";
RL Bur. J. Immunol. 22:2973-2981(1992).
RN [4]
RP SEQUENCE OF 107-192 FROM N.A.
RX MEDLINE=98320644; PubMed=9656488;
RA Harris E.E.; DiStefano T.R.;
RT "Nuclear gene trees and the phylogenetic relationships of the
RT mangabeys (Primates: Papionini).";

```

```

RL Mol. Biol. Evol. 15:892-900(1998).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL, M31134; AAA6838.1; -.
CC EMBL, D63347; BAA09671.1; -.
CC EMBL, X73326; CAA51752.1; -.
CC EMBL, AF057385; AAC25129.1; -.
CC HSPF, P01730; IMR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn.; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin.; ISS.
DR InterPro; IPR008424; CD2.
DR InterPro; IPR009973; CD4_TGAG.
DR InterPro; IPR007110; Ig_V-like.
DR InterPro; IPR003596; Ig_V.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.
DR SMART; SM00406; IGV. 1.
DR PROSITE; PS50835; IGV_LIKE. 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458 T-cell surface glycoprotein CD4.
FT DOMAIN 26 396 Extracellular (Potential).
FT TRANSMEM 397 418 Potential.
FT DOMAIN 419 458 Cytoplasmic (Potential).
FT DOMAIN 26 125 Ig-like V-type.
FT DOMAIN 126 203 Ig-like C2-type 1.
FT DOMAIN 204 317 Ig-like C2-type 2.
FT DOMAIN 318 374 Ig-like C2-type 3.
FT CARBOHYD 296 296 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 325 325 N-linked (GlcNAc...) (By similarity).
FT DISULFID 41 109 By similarity.
FT DISULFID 155 184 By similarity.
FT LIPID 419 419 S-palmitoyl cysteine (By similarity).
FT LIPID 422 422 S-palmitoyl cysteine (By similarity).
FT CONFLICT 42 42 N->T (in Ref. 1).
FT CONFLICT 62 62 L->S (in Ref. 3).
FT CONFLICT 67 67 L->S (in Ref. 2).
FT CONFLICT 169 169 I->L (in Ref. 2).
FT CONFLICT 191 191 K->N (in Ref. 3).
FT CONFLICT 248 248 S->P (in Ref. 3).
FT CONFLICT 265 265 R->Q (in Ref. 3).
FT CONFLICT 349 349 A->T (in Ref. 2).
SQ SEQUENCE 458 AA; 50884 MW; 8BB8039FAFEC808 CRC64;

Query Match 35.4%; Score 945; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 1.9e-51;
Matches 183; Conservative 13; Mismatches 12; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLVLTQALLPATQKGVVLRKSDPTVELTCTASQKSIQPHMKNSNQIK 60
DB 1 NMRGVPFRHLVLTQALLPATQKGVVLRKSDPTVELTCTASQKSIQPHMKNSNQIK 60

```

```

QY 61 ILNQGSLFKGSKLNDRASSRLMDQGNFPLIINKLIEPSDYICEVEDQKEEVQL 120
DB 61 ILGIQSLFKGSKLNDRASSRLMDQGNFPLIINKLIEPSDYICEVENKKEVEL 120
QY 121 LVFGLTANSSTHLLQGSLTLTLSPGSSPSVQCSPPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSSTHLLQGSLTLTLSPGSSPSVQCSPPRGKNIQGGKTLVSQLELDG 180
QY 181 TWTCVTQONOKVKEFKIDIVLAFQKAS 208
DB 181 TWTCVTQONOKVKEFKIDIVLAFQKAS 208

RESULT 6
CD4_MACNE STANDARD; PRT; 458 AA.
AC Q08340; F79196;
DT 01-FEB-1995 (Rel. 31, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Macaca nemestrina (pig-tailed macaque).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Macaca.
OX NCBI_TaxID=9545;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RL Submitted (FEB-1997) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;
RX MEDLINE=93049640; Pubmed=1425921;
RA Fomsgard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus";
RT Eur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63346; BAA09670.1; -
DR EMBL; X73325; CA51751.1; -
DR HSSP; P01730; 1MBR.
DR GO; GO:0042101; C:T-cell receptor complex; ISS.
DR GO; GO:0015026; F:coreceptor activity; ISS.
DR GO; GO:0042289; F:MHC class II protein binding; ISS.
DR GO; GO:0006955; P:immune response; ISS.
DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO; GO:0030217; P:T-cell differentiation; ISS.
DR GO; GO:0045058; P:T-cell selection; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam; PF00047; Ig_2.
DR PRINTS; PR00692; CD4TCANTIGEN.

```

```

DR SMART; SM00406; Ig_v. 1.
DR PROSITE; PS50835; IG_Like; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; Signal; T-cell; Transmembrane.
FT SIGNAL 1 25
FT CHAIN 26 458
FT DOMAIN 26 396
FT TRANSSEM 397 418
FT DOMAIN 419 458
FT DOMAIN 426 458
FT DOMAIN 126 203
FT DOMAIN 204 317
FT DOMAIN 318 374
FT CARBOHYD 42 42
FT CARBOHYD 296 296
FT CARBOHYD 325 325
FT DISULFID 41 109
FT DISULFID 155 184
FT DISULFID 328 370
FT LIPID 419 419
FT LIPID 422 422
FT CONFLICT 57 57
FT CONFLICT 91 91
FT CONFLICT 105 105
FT CONFLICT 113 113
FT CONFLICT 302 302
FT CONFLICT 349 349
SQ SEQUENCE 458 AA; 50905 MW; 751A9BA2CB83EBE6 CRC64;

Query Match 35.4%; Score 944; DB 1; Length 458;
Best Local Similarity 88.0%; Pred. No. 2, 2e-51;
Matches 183; Conservative 14; Mismatches 11; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLTQALPAATQGGKVVYGRKGDVTELTCTSOXKSIQFHHKNSQIK 60
DB 1 MNRGVPFRHLVLTQALPAATQGGKVVYGRKGDVTELTCTSOXKSIQFHHKNSQIK 60
QY 61 ILNQGSLFKGSKLNDRASSRLMDQGNFPLIINKLIEPSDYICEVEDQKEEVQL 120
DB 61 ILGIQSLFKGSKLNDRASSRLMDQGNFPLIINKLIEPSDYICEVENKKEVEL 120
QY 121 LVFGLTANSSTHLLQGSLTLTLSPGSSPSVQCSPPRGKNIQGGKTLVSQLELDG 180
DB 121 LVFGLTANSSTHLLQGSLTLTLSPGSSPSVQCSPPRGKNIQGGKTLVSQLELDG 180
QY 181 TWTCVTQONOKVKEFKIDIVLAFQKAS 208
DB 181 TWTCVTQONOKVKEFKIDIVLAFQKAS 208

RESULT 7
CD4_CERAE STANDARD; PRT; 458 AA.
AC Q08338; Q02805; O77593; Q28217;
DT 01-FEB-1995 (Rel. 31, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 precursor (T-cell surface antigen
T4/Leu-3).
GN Name=CD4;
OS Cercopithecus aethiops (Green monkey) (Griwet).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecinae; Cercopithecus.
OX NCBI_TaxID=9534;
RN [1]
RP SEQUENCE FROM N.A.
RA Hashimoto O., Tatsumi M.;
RT "Molecular cloning and expression of african green monkey CD4.";
RL Submitted (JUL-1996) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE OF 28-424 FROM N.A.
RC TISSUE=Blood;

```

RX MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 RN [3]  
 RP SEQUENCE OF 28-424 FROM N.A.  
 RC TISSUE=Peripheral blood;  
 RA MEDLINE=98017879; PubMed=9379478;  
 RA Fomsgaard A., Mueller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,  
 RA Corbet S., Barre-Sinoussi F., Allan J.S.;  
 RT "Relation between phylogeny of African green monkey CD4 genes and  
 RT their respective simian immunodeficiency virus genes.";  
 RL J. Med. Primatol. 26:120-128(1997).  
 RN [4]  
 RP SEQUENCE OF 107-192 FROM N.A.  
 RA MEDLINE=98320644; PubMed=9656488;  
 RA Harris E.E., Disotell T.R.;  
 RT "Nuclear gene trees and the phylogenetic relationships of the  
 RT mangabeys (Primates: Papionini).";  
 RL Mol. Biol. Evol. 15:892-900(1998).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, D86589; BAA1332.1; -  
 DR EMBL, X73322; CAA51748.1; -  
 DR EMBL, AF001226; AAB60873.1; -  
 DR EMBL, AF001228; AAB60875.1; -  
 DR EMBL, AF057380; AAC25124.1; -  
 DR HSSP; P01730; IMIO.  
 DR GO; GO:0042101; C-T-cell receptor complex; ISS.  
 DR GO; GO:0015026; F:coreceptor activity; ISS.  
 DR GO; GO:0042289; F:MHC class II protein binding; ISS.  
 DR GO; GO:0006955; P:immune response; ISS.  
 DR GO; GO:0045086; P:positive regulation of interleukin-2 biosyn. .; ISS.  
 DR GO; GO:0030217; P:T-cell differentiation; ISS.  
 DR GO; GO:0045058; P:T-cell selection; ISS.  
 DR GO; GO:0007169; P:transmembrane receptor protein tyrosine kin. .; ISS.  
 DR InterPro; IPR008424; CD2.  
 DR InterPro; IPR000973; CD4 TCaG.  
 DR InterPro; IPR007110; Ig-like.  
 DR Pfam; PF00047; Ig\_2.  
 DR PRINTS; PR00632; CD4TCANTIGN.  
 DR PROSITE; PS0035; IG\_LIKE; 1.  
 KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;  
 KW Palmitate; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 25  
 FT CHAIN 26 458  
 FT DOMAIN 26 396  
 FT TRANSMEM 397 418  
 FT DOMAIN 419 458  
 FT DOMAIN 26 125  
 FT DOMAIN 126 203  
 FT DOMAIN 204 317  
 FT DOMAIN 318 374  
 FT CAROHD 42 42  
 FT CAROHD 281 281  
 FT CAROHD 296 296  
 FT CAROHD 325 325  
 FT DISULFID 41 109  
 By similarity.  
 T-cell surface glycoprotein CD4.  
 Extracellular (Potential).  
 Potential.  
 Cytoplasmic (Potential).  
 Ig-like V-type.  
 Ig-like C2-type 1.  
 Ig-like C2-type 2.  
 Ig-like C2-type 3.  
 N-linked (GlcNAc. . .) (Potential).  
 N-linked (GlcNAc. . .) (Potential).  
 N-linked (GlcNAc. . .) (Potential).  
 N-linked (GlcNAc. . .) (Potential).  
 By similarity.

FT DISULFID 155 184 By similarity.  
 FT DISULFID 328 370 By similarity.  
 FT LIPID 419 419 S-palmitoyl cysteine (By similarity).  
 FT LIPID 422 422 S-palmitoyl cysteine (By similarity).  
 FT CONFLICT 46 46 K -> N (in Ref. 2 and 3; AAB60875).  
 FT CONFLICT 59 59 I -> T (in Ref. 3; AAB60873).  
 FT CONFLICT 115 115 K -> E (in Ref. 1).  
 FT CONFLICT 165 165 G -> V (in Ref. 3; AAB60873 and 4).  
 FT CONFLICT 200 200 M -> V (in Ref. 2 and 3).  
 FT CONFLICT 227 227 F -> L (in Ref. 3; AAB60873).  
 FT CONFLICT 271 271 K -> E (in Ref. 3; AAB60873).  
 FT CONFLICT 281 281 N -> H (in Ref. 3; AAB60873).  
 SQ SEQUENCE 458 AA; 51158 MW; FC523D2EDD1F12E7 CRC64;  
 Query Match 34.7%; Score 926; DB 1; Length 458;  
 Best Local Similarity 86.5%; Pred. No. 3e-50;  
 Matches 180; Conservative 12; Mismatches 16; Indels 0; Gaps 0;  
 QY 1 MNRGVPRHLLVLTALPPATOGKRVLTGKGDVETLTASQKSTQFWKSNQIK 60  
 DB 1 NMWGIPEHLLVLTALPPATOGKRVLTGKGDVETLTASQKSTQFWKSNQIK 60  
 QY 61 ILKNGSFLTKGPSKLANDRARSRLMPQGNPPLIKULKTEDSPTYICEVDKEEYOL 120  
 DB 61 ILKNGSFLTKGSSRLRDRSRKSLMDQGSMTIKULKTEDSETTYICEVKNKEEYOL 120  
 QY 121 LVFGLTANSDFHLLQGQSLTTLTSSPGSSPSPVQCRSPRGNKIQGKTLVSQLELDPSG 180  
 DB 121 LVFGLTANSDFHLLQGQSLTTLTSSPGSSPSPVQCRSPRGNKIQGKTLVSQLELDPSG 180  
 QY 181 TWTCTVLOKMKVEFKIDIVLAPQKAS 208  
 DB 181 TWTCTVLOKMKVEFKIDIVLAPQKAS 208  
 QY 191 TWTCTVSDQNTVEFKIDIVLAPQKAS 208  
 DB 191 TWTCTVSDQNTVEFKIDIVLAPQKAS 208  
 RESULT 8  
 CD4\_CERTO  
 ID CD4\_CERTO STANDARD; PRT; 397 AA.  
 AC 008336;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 05-JUL-2004 (Rel. 44, Last annotation update)  
 DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)  
 GN (Fragment).  
 GN Name=CD4;  
 OS Cercopithecus torquatus atys (Red-crowned mangabey) (Sooty mangabey).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;  
 OC Cercopitheciinae; Cercopithecus.  
 OC NCBI\_TaxID=9531;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Blood;  
 RA MEDLINE=93049640; PubMed=1425921;  
 RA Fomsgaard A., Hirsch V.M., Johnson P.R.;  
 RT "Cloning and sequences of primate CD4 molecules: diversity of the  
 RT cellular receptor for simian immunodeficiency virus/human  
 RT immunodeficiency virus.";  
 RL Eur. J. Immunol. 22:2973-2981(1992).  
 CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell  
 CC receptor interaction. May regulate T-cell activation.  
 CC -1- SUBUNIT: Associates with p56-lck (By similarity).  
 CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.  
 CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

```

CC -----
DR EMBL: X73328; CAAS1754.1; -.
DR EMBL: X73327; CAAS1753.1; -.
DR HSSP: P01730; 1MIQ.
DR GO: GO:0042101; C/T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR005596; Ig_v.
DR Pfam: PF00047; Ig; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; T-cell; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 369 Extracellular (Potential).
FT TRANSMEM 370 391 Potential.
FT DOMAIN 392 >397 Cytoplasmic (Potential).
FT DOMAIN <1 98 Ig-like V-type.
FT DOMAIN 99 176 Ig-like C2-type 1.
FT DOMAIN 177 290 Ig-like C2-type 2.
FT DOMAIN 291 347 Ig-like C2-type 3.
FT CARBOHYD 15 15 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (Potential).
FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (Potential).
FT DISULFID 14 82 By similarity.
FT DISULFID 128 157 By similarity.
FT DISULFID 301 343 S-palmitoyl cysteine (By similarity).
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 Missing.
FT VARIANT 20 20 T -> I.
FT VARIANT 43 43 N -> D.
FT VARIANT 86 86 F -> L.
FT VARIANT 96 96 V -> M.
FT VARIANT 173 173 V -> K.
FT VARIANT 316 316 R -> K.
FT NON_TER 397
SQ SEQUENCE 397 AA; 43926 MW; 86608636D2DB38A7 CRC64;

Query Match 30.7%; Score 820; DB 1; Length 397;
Best Local Similarity 87.8%; Pred. No. 1.2e-43;
Matches 159; Conservative 11; Mismatches 11; Indels 0; Gaps 0;

```

```

DT 01-FEB-1995 (Rel. 31, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE T-cell surface glycoprotein CD4 (T-cell surface antigen T4/Leu-3)
DE (Fragment).
GN Name=CD4.
OS Erythrocybus patas (Red guenon) (Cercopithecus patas).
OC Eukaryota; Metazoa; Chordata; Carnate; Vertebrata; Euarchontomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
OC Cercopithecoidea; Erythrocybus.
OX NCBI_TaxID=9538;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Blood;
RA MEDLINE=93049640; PubMed=1425921;
RA Fomsgaard A., Hirsch V.M., Johnson P.R.;
RT "Cloning and sequences of primate CD4 molecules: diversity of the
RT cellular receptor for simian immunodeficiency virus/human
RT immunodeficiency virus.";
RL Bur. J. Immunol. 22:2973-2981(1992).
CC -1- FUNCTION: Accessory protein for MHC class-II antigen/T-cell
CC receptor interaction. May regulate T-cell activation.
CC -1- SUBUNIT: Associates with p56-lck (By similarity).
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- SIMILARITY: Contains 3 immunoglobulin-like C2-type domains.
CC -1- SIMILARITY: Contains 1 immunoglobulin-like V-type domain.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X73324; CAAS1750.1; -.
DR HSSP: P01730; 1MIQ.
DR GO: GO:0042101; C/T-cell receptor complex; ISS.
DR GO: GO:0015026; F:coreceptor activity; ISS.
DR GO: GO:0042289; F:MHC class II protein binding; ISS.
DR GO: GO:0006955; P:immune response; ISS.
DR GO: GO:0045086; P:positive regulation of interleukin-2 biosyn. . .; ISS.
DR GO: GO:0030217; P:T-cell differentiation; ISS.
DR GO: GO:0045058; P:T-cell selection; ISS.
DR GO: GO:0007169; P:transmembrane receptor protein tyrosine kin. . .; ISS.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TCAG.
DR InterPro: IPR007110; IG-like.
DR InterPro: IPR003596; IGV_v.
DR Pfam: PF00047; Ig; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SM00406; IGV; 1.
DR PROSITE: PSS0835; IG LIKE; 1.
KW Glycoprotein; Immune response; Immunoglobulin domain; Lipoprotein;
KW Palmitate; Repeat; T-cell; Transmembrane.
FT NON_TER 1
FT DOMAIN <1 369 Extracellular (Potential).
FT TRANSMEM 370 391 Potential.
FT TRANSMEM 392 >397 Cytoplasmic (Potential).
FT DOMAIN <1 98 Ig-like V-type.
FT DOMAIN 99 176 Ig-like C2-type 1.
FT DOMAIN 177 290 Ig-like C2-type 2.
FT DOMAIN 291 347 Ig-like C2-type 3.
FT CARBOHYD 269 269 N-linked (GlcNAc. . .) (By similarity).
FT CARBOHYD 298 298 N-linked (GlcNAc. . .) (By similarity).
FT DISULFID 14 82 By similarity.
FT DISULFID 128 157 By similarity.
FT DISULFID 301 343 By similarity.
FT LIPID 392 392 S-palmitoyl cysteine (By similarity).
FT LIPID 395 395 S-palmitoyl cysteine (By similarity).
FT NON_TER 397
SQ SEQUENCE 397 AA; 44081 MW; 67887397A6B7E4AF CRC64;

Query Match 30.7%; Score 819; DB 1; Length 397;

```

Best Local Similarity 87.3%; Pred. No. 1.4e-43;  
Matches 158; Conservative 10; Mismatches 13; Indels 0; Gaps 0;

Qy	28	VLAGGGTVELTCTASOKKSI OFHKNKSNOKI LKGNOS FLTKGSKLNDRADSRSLW	87
Qy		1	VVLKGGGTVELTCNASQKTTTOFHKNKSNOMKILKQOS FLTKGSKLDRDTSRKSJW
Db			60
Qy	88	DQGNPELIIKNLIKIEDSDTYICEVEBOKEEVOLVLEGLTANSPDTHLLOQSLTLTESPP	147
Qy		61	DQGCSTMIHKLIKIEDSEFYICEVEBOKKEVELVLEGLTANSPDTHLLOQSLTLTESPP
Db			120
Qy	148	GSSPSEVQCRSPKGNKIQGSKTLSVSOLELODSGTWCTVYLOQKXVEPKIDIVIAFOXA	207
Db	121	GSSPSEVQCRSPKGNKIQGRTLSVPLERODSDGTWCTVSOQDNTVEPKIDIVIAFOXA	180
Qy		208 S 208	
Db		181 S 181	

RESULT 10

ID	009261	PRELIMINARY;	PRT;	397 AA.
AC	009261;			
DT	01-JUL-1997	(TREMBLrel. 04, Created)		
DT	01-JUL-1997	(TREMBLrel. 04, Last sequence update)		
DT	01-OCT-2003	(TREMBLrel. 25, Last annotation update)		
DE	CD4 (Fragment) .			
GN	Name=CD4;			
OS	Cercopithecus sabaues.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;			
OC	Cercopitheciinae; Cercopithecus.			
OX	NCBI_TaxID=60711;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98017879; PubMed=9379478;			
RA	Fombes A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,			
RA	Cotbet S., Baire-Sinoussi F., Allan J.S.;			
RT	"Relation between phylogeny of African green monkey CD4 genes and			
RT	their respective simian immunodeficiency virus genes."			
RL	J. Med. Primatol. 26:120-128(1997).			
DR	EMBL: AF001225; AAB60872.1; -			
DR	GO: GO:0016021; C:integral to membrane; IEA.			
DR	GO: GO:0007155; P:cell adhesion; IEA.			
DR	GO: GO:0006955; P:immune response; IEA.			
DR	InterPro: IPR008424; CD2.			
DR	InterPro: IPR000973; CD4 TCAG.			
DR	InterPro: IPR007110; Ig-like.			
DR	InterPro: IPR003596; Ig v.			
DR	Pfam: PF05790; C2-sec; 2.			
DR	Pfam: PF00047; Ig; 1.			
DR	PRINTS; PR00692; CD4TCAGTIGEN.			
DR	SMART; SM00406; IGv; 1.			
DR	PROSITE; PSS0835; IG_LIKE; 1.			
FT	NON_TER	1		
FT	NON_TER	397	397	
Q	SEQUENCE	397 AA; 43345 MW; 80C963B92A868CD3 CRC64;		

	Query March	30.1%;	Score 803;	DB 2;	Length 397;	
	Best Local Similarity	86.2%;	Pred. No. 1.4e-42;			
	Matches 156;	Conservative 10;	Mismatches 15;	Indels 0;	Gaps 0	
QY	28	VTLGKGGTVLTLCTASOKKSIOFHWNKSNQIKIINGQSFLTKGPSKLNDADSRSLW	87			
		:	:			
Db	1	VVLAKGDTVELTCNASONTTTFHMKNSNQIKIGKQSSPLTKGSCKLRIDSRKSLW	60			
		:	:			
QY	88	DGNFPELLIKULIEDSDPTICEVDQKEEVLVFGLTANDTHLQGOSLTLESP	147			
		:	:			
Db	61	DQGFSMITIKULKIEDSETYICEVNEKEEVLLVFGLTANDTHLQGOSTLLESP	120			
		:	:			
QY	148	GSSPSVGCRSPRKNIQGGKTLSVQLDELQDSGTCTYLQNKVEFKLDIVLAFQA	207			
		:	:			

Db 121 GSSPSVKCRSPGKNIQGRTLSVPQLERQDSGTWTCTVSQDQNMVEFKIDIVLAFQKA 180

QY	208 S	208
Db	181 S	181

## RESULT 11

ID	009262;	PRELIMINARY;	PRT;	397 AA.
AC				
DT	01-JUL-1997 (TrEMBLrel. 04, Created)			
DT	01-JUL-1997 (TrEMBLrel. 04, Last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)			
DE	CD4 (Fragment).			
GN	Name=CD4;			
OS	Cercopithecus tantalus.			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;			
OC	Cercopithecinae; Cercopithecus.			
OX	NCBI_TaxID=60712;			
RN	[1]			
RP	SEQUENCE FROM N. A.			
RX	MEDLINE=98017879; PubMed=9379478;			
RA	Fomsgaard S., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.			
RA	Corbet S., Barre-Sinoussi F., Allan J.S.;			
RT	"Relation between phylogeny of African green monkey CD4 genes and			
RT	their respective simian immunodeficiency virus genes."			
RL	J. Med. Primatol. 26:120-128(1997).			
DR	EMBL, AF001421; AAB60868.1; --			
DR	GO; GO:0016021; C:Integral to membrane; IEA.			
DR	GO; GO:0007155; P:cell adhesion; IEA.			
DR	GO; GO:0006955; P:immune response; IEA.			
DR	InterPro; IPR008424; CD2.			
DR	InterPro; IPR009733; CD4_TCAG.			
DR	InterPro; IPR007110; Ig-Like.			
DR	InterPro; IPR003596; Ig_V.			
DR	Pfam; PF05790; C2-set; 2.			
DR	Pfam; PF00047; Ig; 1.			
DR	PRINTS; PR00692; CD4TCANTIGEN.			
DR	SMART; SMO0406; IGV; 1.			
DR	PROSITE; PS50835; IG_Like; 1.			
FT	NON_TER	1	_1	
FT	NON_TER	397		
QC	SEQUENCE	397 AA;	43954 MW;	CF7E2F5D82335B0D CRC64;

Query Match	30.1%;	Score 803;	DB 2;	Length 397;
Best Local Similarity	86.2%;	Pred. No. 1.4e-42;		
Matches 156;	Conservative 10;	Mismatches 15;	Indels 0;	Gaps 0

QY 26 VVLGGKGGTVELTCTASOKKSIOFHMKNSNOIKILNQGSFLFKGSKLNDRADRSRLW 87

Db 1 VVLGGKGGTVELTCTASONTTTOFHMKNSNOIKILNQGSFLFKGSSKLRDRIDSKSLW 60

QY 88 DQGNFPLIIRKULKIEDSPDYICEVEDQKEQYLVFGLTANSTDTLLQOQSITLTLESPP 147

Db 61 DQGSCTSMIIRKULKIEDSEYIICEVENKKEEVELLVGLTANSTDTLLQOQSITLTLESPP 120

QY 148 GSSPSEVQCRSPRGNKNIQSGKTLISQLELDDSGMTCTYLQONKTYEPIKIDIVVLVAFQA 207

Db 121 GSSPSEVQCRSPRGNKNIQSGKTLISQLELDDSGMTCTYLQONKTYEPIKIDIVVLVAFQA 180

QY	208	S	208
Db	181	S	181

RESULT 12	
009263	
ID 009263	PRELIMINARY;
AC 009263:	PRT; 397 AA.
DT 01-JUL-1997	(TREMBlrel. 04, Created)
PT 01-JUL-1997	(TREMBlrel. 04, last sequence update)

```

DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus tantalus.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_Taxid=60712;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001222; AAB60869.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SMO0406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT SEQUENCE 397 AA; 43994 MW; A3CD031535A51524 CRC64;
SQ
Query Match 30.1%; Score 803; DB 2; Length 397;
Best Local Similarity 86.2%; Pred. No. 1.4e-42;
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VILGKGDVLTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGPKLNDRADSRSLW 87
DB 1 VILGKGDVLTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGPKLNDRADSRSLW 60

QY 88 DQGNFPLIINKLIEDSDTYICEVEDQKEVQLVGLTANSTHLLQGSLTLTLESPP 147
DB 61 DQGNFPLIINKLIEDSEITYICEVENKKEVELLVGLTANSTHLLQGSLTLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLISVSOLELDSGTWTCTVLOKKEVFKIDIIVLAFQKA 207
DB 121 GSSPSVQCRSPRGKNIQGGKTLISVPOLEKDSGTWTCTVSDQNTVEFKIDIIVLAFQKA 180

QY 208 S 208
DB 181 S 181

RESULT 13
Q95NBE9 PRELIMINARY; PRT; 397 AA.
ID 095NBE9;
AC 095NBE9;
DT 01-DEC-2001 (TReMBLrel. 19, Created)
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus pygerythrus (Vervet monkey).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_Taxid=60710;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;

```

```

RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001227; AAB60874.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.
DR PRINTS: PR00692; CD4TCANTIGEN.
DR SMART: SMO0406; IGV; 1.
DR PROSITE: PSS0835; IG_LIKE; 1.
FT NON_TER 1
FT SEQUENCE 397 AA; 43946 MW; 21C3E30882ABFBC0 CRC64;
SQ
Query Match 30.1%; Score 803; DB 2; Length 397;
Best Local Similarity 86.2%; Pred. No. 1.4e-42;
Matches 156; Conservative 10; Mismatches 15; Indels 0; Gaps 0;

QY 28 VILGKGDVLTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGPKLNDRADSRSLW 87
DB 1 VILGKGDVLTCTASQKKSIOFHWNKSNQIKILGNQGSFLTKGPKLNDRADSRSLW 60

QY 88 DQGNFPLIINKLIEDSDTYICEVEDQKEVQLVGLTANSTHLLQGSLTLTLESPP 147
DB 61 DQGNFPLIINKLIEDSEITYICEVENKKEVELLVGLTANSTHLLQGSLTLTLESPP 120

QY 148 GSSPSVQCRSPRGKNIQGGKTLISVSOLELDSGTWTCTVLOKKEVFKIDIIVLAFQKA 207
DB 121 GSSPSVQCRSPRGKNIQGGKTLISVPOLEKDSGTWTCTVSDQNTVEFKIDIIVLAFQKA 180

QY 208 S 208
DB 181 S 181

RESULT 14
O09259 PRELIMINARY; PRT; 397 AA.
ID 009259;
AC 009259;
DT 01-JUL-1997 (TReMBLrel. 04, Created)
DT 01-JUL-1997 (TReMBLrel. 04, Last sequence update)
DT 01-OCT-2003 (TReMBLrel. 25, Last annotation update)
DE CD4 (Fragment) .
GN Name=CD4;
OS Cercopithecus sabaues.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Cercopitheciidae;
OC Cercopithecinae; Cercopithecus.
NCBI_Taxid=60711;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=98017879; PubMed=9379478;
RA Fomsgaard A., Muller-Trutwin M.C., Diop O., Hansen J., Mathiot C.,
RA Corbet S., Barre-Sinoussi F., Allan J.S.;
RT "Relation between phylogeny of African green monkey CD4 genes and
RT their respective simian immunodeficiency virus genes.";
RL J. Med. Primatol. 26:120-128(1997).
DR EMBL: AF001223; AAB60870.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0007155; P:cell adhesion; IEA.
DR GO: GO:0006955; P:immune response; IEA.
DR InterPro: IPR008424; CD2.
DR InterPro: IPR000973; CD4_TcAg.
DR InterPro: IPR007110; Ig-Like.
DR InterPro: IPR003596; Ig_v.
DR Pfam: PF05790; C2-set; 2.

```



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 22, 2004, 01:20:54 : Search time 6279 Seconds  
(without alignments)  
10845.225 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgacacgggggggtccctt.....gtcagcagatcatactacg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 2364849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : GenEmbl:\*  
1: gb\_ba:\*  
2: gb\_hg:\*  
3: gb\_in:\*  
4: gb\_om:\*  
5: gb\_ov:\*  
6: gb\_pat:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sls:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1440	100.0	1440	6	BD268881 Novel chi
2	625	43.4	1213	6	AR363716 Sequence
3	624.4	43.4	1377	6	AX100880 Sequence
4	624.4	43.4	1742	6	AR380468 Sequence
5	624.4	43.4	1742	6	AX287109 Sequence
6	624.4	43.4	1742	9	HUMATCP4
7	624.4	43.4	1742	9	HUMATCP4A
8	624.4	43.4	1910	6	I08116
9	624.4	43.4	3133	6	I08115
10	624.4	43.4	7533	12	AY438650
11	622.8	43.2	1273	6	I06223
12	622.8	43.2	1273	6	I07147
13	622.8	43.2	1304	6	AR067943
14	622.8	43.2	1304	6	AR166802
15	622.8	43.2	1389	6	AR062468
16	622.8	43.2	1389	6	AR067924
17	622.8	43.2	1389	6	AR166783
18	622.8	43.2	1537	6	C0722428
19	622.8	43.2	1599	6	AR062469

20	622.8	43.2	1599	6	AR067925
21	622.8	43.2	1599	6	AR166784
22	622.8	43.2	1728	6	AR062467
23	622.8	43.2	1728	6	AR067923
24	622.8	43.2	1728	6	AR166782
25	622.8	43.2	1742	6	AR104143
26	622.8	43.2	1742	6	I08673
27	622.8	43.2	1742	6	AR339862
28	622.8	43.2	1742	6	AR448985
29	622.8	43.2	2589	6	AR392157
30	622.8	43.2	3084	9	BC025782
31	621.2	43.1	1273	6	AR104142
32	621.2	43.1	1273	6	AR448979
33	621.2	43.1	1416	6	I07030
34	621.2	43.1	1416	6	I09301
35	621.2	43.1	1416	6	I27613
36	621.2	43.1	1416	6	AR489592
37	621.2	43.1	1742	6	I07208
38	621.2	43.1	2940	6	AR064430
39	621.2	43.1	3064	6	I09237
40	621.2	43.1	3426	6	AR064429
41	621.2	43.1	3480	6	AR064431
42	621.2	43.1	3721	6	AR064432
43	621.2	43.1	3993	6	AR064433
44	621.2	43.1	4059	6	AR064434
45	621.2	43.1	4632	6	AR064435

## ALIGNMENTS

RESULT 1	BD268881	1440 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD268881				
DEFINITION	Novel chimeric protein for prevention and treatment of HIV infection.				
ACCESSION	BD268881.1	GI:33078649			
VERSION	BD268881				
KEYWORDS	JP 2002538814-A/1.				
SOURCE	JP 2002538814-A/1.				
ORGANISM	synthetic construct				
REFERENCE	1 (bases 1 to 1440)				
AUTHORS	Berger, E.A. and Castillo, C.D.				
TITLE	Novel chimeric protein for prevention and treatment of HIV infection				
JOURNAL	Patent: JP 2002538814-A 1 19-NOV-2002;				
	THE GOVERNMENT OF THE UNITED STATES OF AMERICA AS REPRESENTED BY				
	THE AUSTRIAN NORIC BIOTHERAPEUTICS AKTIENGESELLSCHAFT SECRETARY				
	DEPARTMENT OF HEALTH AND HUMAN SERVICES THE NATIONAL INSTITUTES OF				
	HEALTH				
OS	Artificial Sequence				
PN	JP 2002538814-A/1				
PD	19-NOV-2002				
PR	16-MAR-2000 JP 2000605633				
PT	EDWARD A BERGER, CHRISTIE DEL CASTILLO				
PC	C12N15/09, A61K38/00, A61P31/18, C07K5/103, C07K14/00, C07K14/155,				
PC	C07K14/715,				
PC	C07K14/73, C07K16/10, C07K19/00, C12N5/10, C12P21/02, C12P21/02,				
PC	C12R1/91				
PC	C12N15/00, C12N5/00, A61K37/02				
CC	Description of Artificial Sequence: CD4-scFv(17b) FH				
FT	Location/Qualifiers				
FT	1..1440				
FEATURES	Location/Qualifiers				
source	1..1440				
	/organism="Artificial Sequence".				
	/organism="synthetic construct"				
	/mol_type="genomic DNA"				
	/db_xref="taxon:32630"				
ORIGIN					

Query Match 100.0%; Score 1440; DB 6; Length 1440;  
 Best Local Similarity 100.0%; Pred. No. 0;  
 Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 QY 61 GCAGCCACTCAGGGGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 61 GCAGCCACTCAGGGGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 QY 121 TGTACAGCTTCCGAAAGAAAGCATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 DB 121 TGTACAGCTTCCGAAAGAAAGCATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 QY 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGGTCCATCCAGCTGAATGATCGGCT 240  
 DB 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGGTCCATCCAGCTGAATGATCGGCT 240  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCGCTGATCAAGAACTTAAG 300  
 DB 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCGCTGATCAAGAACTTAAG 300  
 QY 301 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGCAAGAAAGAGAGTGCATTTG 360  
 DB 301 ATGAAGACTCAGATTAATCATCTGTGAAGTGAAGCAAGAAAGAGAGTGCATTTG 360  
 QY 361 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTACC 420  
 DB 361 CTAGTGTTCGATTTGACTGCCAATCTGACACCACTGCTTCAAGGGGAGAGCTTACC 420  
 QY 421 CTGACCTTGAAGAGCCCCCTGTGAGTGAAGCCCTCAGTGCATGTAAGAGTCAAGGGGT 480  
 DB 421 CTGACCTTGAAGAGCCCCCTGTGAGTGAAGCCCTCAGTGCATGTAAGAGTCAAGGGGT 480  
 QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGACTCAGATAGTGGC 540  
 DB 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGACTCAGATAGTGGC 540  
 QY 541 ACCTGGAATGCACTGTCTTTCAGAACCAAGAAAGTGAAGTTCAAATAGACATGTG 600  
 DB 541 ACCTGGAATGCACTGTCTTTCAGAACCAAGAAAGTGAAGTTCAAATAGACATGTG 600  
 QY 601 GTGCTACCTTCCAGAGAGGCTCCGAGAGTGGCGGTGTAAGGGGAGGCTTCAAGGGCA 660  
 DB 601 GTGCTACCTTCCAGAGAGGCTCCGAGAGTGGCGGTGTAAGGGGAGGCTTCAAGGGCA 660  
 QY 661 GGTGATCCGGTGGCGAGGCTCGGCGGGGTGGAACCGGGGCTGGCGCTCCGAGGC 720  
 DB 661 GGTGATCCGGTGGCGAGGCTCGGCGGGGTGGAACCGGGGCTGGCGCTCCGAGGC 720  
 QY 721 GAGAGTTCACAGTGTGCTGTGAGTCTGGGCTGAAGTGAAGACCTGGCTCTCG 780  
 DB 721 GAGAGTTCACAGTGTGCTGTGAGTCTGGGCTGAAGTGAAGACCTGGCTCTCG 780  
 QY 781 GTGAAGGTCTCCGCAAGGCTCTGAGAGCACTTCACTCAATATATAGTTTAACTGGG 840  
 DB 781 GTGAAGGTCTCCGCAAGGCTCTGAGAGCACTTCACTCAATATATAGTTTAACTGGG 840  
 QY 841 CGACAGGCTCTGAGCAAGGCTTGAAGTGAAGAGATCATCACTATCTTGATGA 900  
 DB 841 CGACAGGCTCTGAGCAAGGCTTGAAGTGAAGAGATCATCACTATCTTGATGA 900  
 QY 901 GCACACTACGACCGCACTTCAGAGGCAAGTCAAGTTACCGGGCAAGTCCAGAGC 960  
 DB 901 GCACACTACGACCGCACTTCAGAGGCAAGTCAAGTTACCGGGCAAGTCCAGAGC 960  
 QY 961 ACAGTCACTGAGAGCGGCAATCTAAGATCTGACATACAGAGCGGATATTTCTGGG 1020  
 DB 961 ACAGTCACTGAGAGCGGCAATCTAAGATCTGACATACAGAGCGGATATTTCTGGG 1020  
 QY 1021 GAGTGTACGAGGAGAGCGGACGAGGGGGAATATGATTAATATGGTTTCTGAACAT 1080

DB 1021 GAGTGTACGAGGAGAGCGGACGAGGGGGAATATGATTAATATGGTTTCTGAACAT 1080  
 QY 1081 TGGGGCCAGGAAACCTTGCTCAGGTCACTCAGTGTGGCGGTCCCGAAGTGTGGG 1140  
 DB 1081 TGGGGCCAGGAAACCTTGCTCAGGTCACTCAGTGTGGCGGTCCCGAAGTGTGGG 1140  
 QY 1141 AGCGGTGGCGGATCTGAACTCGAGTTGACGCACTCCAGCCACCTGTCTGTCT 1200  
 DB 1141 AGCGGTGGCGGATCTGAACTCGAGTTGACGCACTCCAGCCACCTGTCTGTCT 1200  
 QY 1201 CCAGGGAAAGACCACTCTCTTCAGAGGCCAGTGAAGTGTATGACGACTTAAGCC 1260  
 DB 1201 CCAGGGAAAGACCACTCTCTTCAGAGGCCAGTGAAGTGTATGACGACTTAAGCC 1260  
 QY 1261 TGGTACCAAGAAACCTGGCCAGGCTCCAGGCTCTCTCAATATATGGTATCAACAGG 1320  
 DB 1261 TGGTACCAAGAAACCTGGCCAGGCTCCAGGCTCTCTCAATATATGGTATCAACAGG 1320  
 QY 1321 GCCACCGGTGCCAGCAGATTCAATGAGAGTGGGTCTGGGACAAATCACTTCACC 1380  
 DB 1321 GCCACCGGTGCCAGCAGATTCAATGAGAGTGGGTCTGGGACAAATCACTTCACC 1380  
 QY 1381 ATCAGCAGCTGCAATCTGAAGATTTTGCAGTTTACTGTGACAGTACAACTGG 1440  
 DB 1381 ATCAGCAGCTGCAATCTGAAGATTTTGCAGTTTACTGTGACAGTACAACTGG 1440

RESULT 2  
 AR363716 1213 bp DNA linear PAT 03-SEP-2003  
 LOCUS AR363716  
 DEFINITION Sequence 1 from patent US 5223418.  
 ACCESSION AR363716  
 VERSION AR363716.1 GI:34425655  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unclassified.  
 REFERENCE 1 (bases 1 to 1213)  
 AUTHORS Arcuri,E.J., Brawner,M.E., Donovan,M.J., Gerber,R.G. and Keller,J.A.  
 TITLE Method of Improving the yield of heterologous proteins produced by Streptomyces lividans  
 JOURNAL Patent: US 5223418-A 1 29-JUN-1993;  
 FEATURES location/Qualifiers  
 source 1..1213  
 /organism="unknown"  
 /mol\_type="genomic DNA"

ORIGIN  
 Query Match 43.4%; Score 625; DB 6; Length 1213;  
 Best Local Similarity 97.7%; Pred. No. 1.3e-134;  
 Matches 634; Conservative 0; Mismatches 15; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 79 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 138  
 QY 61 GCAGCCACTCAGGGGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 120  
 DB 139 GCAGCCACTCAGGGGAAAGAGTGTGTGGGCAAAAAGGGGATACAGTGAACCTGACC 198  
 QY 121 TGTACAGCTTCCGAAAGAAAGCATCAATTCACCTGGAATAAACTCCACAGATAAG 180  
 DB 199 TGTACAGCTTCCGAAAGAAAGCATCAATTCACCTGGAATAAACTCCACAGATAAG 258  
 QY 181 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGGTCCATCCAGCTGAATGATCGGCT 240  
 DB 259 ATTCTGGGAATCAGGGCTCTTTTAACTTAAGGTCCATCCAGCTGAATGATCGGCT 318  
 QY 241 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCGCTGATCAAGAAATCTTAAG 300  
 DB 319 GACTCAAGAAAGACCTTTGGGACCAAGGAACTCCCGCTGATCAAGAAATCTTAAG 378

QY 301 ATGAGAGCTCAGATCTTACATCTGTGAGTGAAGAGACACAGAGGAGGAGGTCGAATTG 360  
 DB 379 ATGAGAGCTCAGATCTTACATCTGTGAGTGAAGAGACACAGAGGAGGAGGTCGAATTG 438  
 QY 361 CTAGTGTTCGAGTTGACTGCACTCTGACACCCACCTGCTTCAGAGGAGAGGCTGACC 420  
 DB 439 CTAGTGTTCGAGTTGACTGCACTCTGACACCCACCTGCTTCAGAGGAGAGGCTGACC 498  
 QY 421 CTGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 DB 499 CTGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 558  
 QY 481 AAAAATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 DB 559 AAAAATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 618  
 QY 541 ACCTGAGACATGACCTGTCTTGGAGAACACAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 DB 619 ACCTGAGACATGACCTGTCTTGGAGAACACAGAGAGGAGGAGGAGGAGGAGGAGGAGG 678  
 QY 601 GTGCTAGCTTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 649  
 DB 679 GTGCTAGCTTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 727

RESULT 3  
 AX100880 1377 bp DNA linear PAT 10-APR-2001  
 LOCUS Sequence 30 from Patent WO0122084.  
 DEFINITION AX100880  
 ACCESSION AX100880  
 VERSION AX100880.1 GI:13619792  
 KEYWORDS  
 ORGANISM Homo sapiens (human)  
 SOURCE Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.

REFERENCE 1  
 Jaksoben, B.K.  
 TITLE Spr identification of inhibitors of receptor-ligand interactions  
 JOURNAL Patent: WO 0122084-A 30 29-MAR-2001;  
 Avidec Ltd (GB)

FEATURES  
 source Location/Qualifiers

## CDS

1..1377  
 /organism="Homo sapiens"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:9606"  
 1..1377  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAC36495.1"  
 /db\_xref="GI:13619793"  
 /translation="MNRGVPRFHLVLLVQLALLPATQKVVLGKGPVVELNCTAS  
 OKSIQFHKNSNOIKILNGOSFLTKGPSKLNDASRSRSLMDGNPLIKNLKIE  
 DSDYICVEDQKEVOLVPLGLTANSTHILIGOSLITLSPSSSYVCKESPRG  
 KNIQGGKTLSSQLELDQSGTTCVYLNOKKVEKRIIVLAFQKASISYVKKGEQ  
 VEFSPFLAVTEKLTVSGELMWQARASSSKWTFDLKNEVSVKRVTVQDPKIQMGK  
 KLPLTLTQALPOYAGSGNLTALVANTGKLHGVNLYVWRATOLQNLTCFVWGP  
 SPKMLSLKLEKKAQVSKREKAVVWNPENAGMQLSDSGVLLSNIKVLPWST  
 PVQPMALIVLGVAGLILFIGLIFPCVRCRRRRQAEMSGIQRLLSEKTKCQCPHR  
 FQKTCSPF"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1377;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 QY 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120  
 DB 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120

DB 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120  
 QY 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAAATCCACAGATTAAG 180  
 DB 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAAATCCACAGATTAAG 180  
 QY 181 ATTCTGGGAAATCAGAGGCTCTCTTTAACTAAAGGTCATCCAGCTGAATGATGCGCT 240  
 DB 181 ATTCTGGGAAATCAGAGGCTCTCTTTAACTAAAGGTCATCCAGCTGAATGATGCGCT 240  
 QY 241 GACTCAGAGAGAGGCTTTGGGACCAAGAAATTTCCCTGTATCATCAAGATCTTAAG 300  
 DB 241 GACTCAGAGAGAGGCTTTGGGACCAAGAAATTTCCCTGTATCATCAAGATCTTAAG 300  
 QY 301 ATGAGAGCTCAGATCTTACATCTGTGAGTGAAGAGACACAGAGAGGAGGTCGAATTG 360  
 DB 301 ATGAGAGCTCAGATCTTACATCTGTGAGTGAAGAGACACAGAGAGGAGGTCGAATTG 360  
 QY 361 CTAGTGTTCGAGTTGACTGCACTCTGACACCCACCTGCTTCAGAGGAGAGGCTGACC 420  
 DB 361 CTAGTGTTCGAGTTGACTGCACTCTGACACCCACCTGCTTCAGAGGAGAGGCTGACC 420  
 QY 421 CTGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 DB 421 CTGACCTTGAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 480  
 QY 481 AAAAATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 DB 481 AAAAATCATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 540  
 QY 541 ACCTGAGACATGACCTGTCTTGGAGAACACAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 DB 541 ACCTGAGACATGACCTGTCTTGGAGAACACAGAGAGGAGGAGGAGGAGGAGGAGGAGG 600  
 QY 601 GTGCTAGCTTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 626  
 DB 601 GTGCTAGCTTTCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 626

RESULT 4  
 AR380468 1742 bp DNA linear PAT 18-DEC-2003  
 LOCUS Sequence 1013 from patent US 6607879.  
 DEFINITION AR380468  
 ACCESSION AR380468  
 VERSION AR380468.1 GI:40088102  
 KEYWORDS  
 SOURCE Unknown.  
 ORGANISM Unknown.  
 Unclassified.

REFERENCE 1 (bases 1 to 1742)  
 Cocks, B.G., Stuart, S.G. and Seilhamer, J.J.  
 TITLE Compositions for the detection of blood cell and immunological  
 response gene expression  
 JOURNAL Patent: US 6607879-A 1013 19-AUG-2003;  
 location/Qualifiers

FEATURES  
 source

1..1742  
 /organism="unknown"  
 /mol\_type="genomic DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1742;  
 Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 60  
 DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCACTGGCGCTCTCCCA 135  
 QY 61 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 120  
 DB 136 GCAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAGGGGATACAGTGAACGTACC 195  
 QY 121 TGTACAGCTTCCCAAGAGAGCATACATTTCCACTGGAAAAATCCACAGATTAAG 180

```

Db      196 TGTCACCTTCCAGAAAGACATACATTCACCTGGAAAAAATCCAAACCAATAAAG 255
Qy      181 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAAGTCATCCAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAAGTCATCCAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 375
Qy      301 ATGAGAGCTCAGATCTTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 360
Db      376 ATGAGAGCTCAGATCTTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 435
Qy      361 CTAGTGTGGGATGACTGTGCACTGCAACCACTGCTTCAAGGGGACAGACCTGACC 420
Db      436 CTAGTGTGGGATGACTGTGCACTGCAACCACTGCTTCAAGGGGACAGACCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 555
Qy      481 AAAAACTACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 540
Db      556 AAAAACTACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 615
Qy      541 ACCTGGACATGCACTGTCTTGGCAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600
Db      616 ACCTGGACATGCACTGTCTTGGCAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 675
Qy      601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

```

```

RESULT 5
AX287109      1742 bp      DNA      linear      PAT 21-NOV-2001
LOCUS      AX287109
DEFINITION      Sequence 7 from Patent WO0164752.
ACCESSION      AX287109
VERSION      AX287109.1 GI:17049085
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS      Liltman,D.R., Kwon,D., van Kooyk,Y.C. and Geijtenbeek,T.C.
TITLES      Antibody inhibiting the binding between gp120 and dc-si3n and
screening methods
JOURNAL      Patent: WO 0164752-A 7 07-SEP-2001;
NEW YORK UNIVERSITY (US); KATHOLIEKE UNIVERSITEIT NIJMEGEN (NL)
FEATURES
source
1..1742
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

```

```

ORIGIN
Query Match      43.4%; Score 624.4; DB 6; Length 1742;
Best Local Similarity 99.8%; Pred.No.1.7e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy      1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60
Db      76 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 135
Qy      61 GCACCACTCAGGAGAAAGAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 120
Db      136 GCACCACTCAGGAGAAAGAGTGTGTGTGGCAAAAAAGGGGATACAGTGAATGACC 195
Qy      121 TGTCAGCTTCCAGAAAGACATACATTCACCTGGAAAAAATCCAAACCAATAAAG 180

```

```

Db      196 TGTCACCTTCCAGAAAGACATACATTCACCTGGAAAAAATCCAAACCAATAAAG 255
Qy      181 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAAGTCATCCAGCTGAATGATCGGCT 240
Db      256 ATTCGGGAATCAGGGCTCCTTCTTAACCTAAAGTCATCCAGCTGAATGATCGGCT 315
Qy      241 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 300
Db      316 GACTCAAGAAAGACCTTTGGGACCAAGAACTTCCCTGATCATCAAGAAATCTTAAG 375
Qy      301 ATGAGAGCTCAGATCTTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 360
Db      376 ATGAGAGCTCAGATCTTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGAGTCAATTG 435
Qy      361 CTAGTGTGGGATGACTGTGCACTGCAACCACTGCTTCAAGGGGACAGACCTGACC 420
Db      436 CTAGTGTGGGATGACTGTGCACTGCAACCACTGCTTCAAGGGGACAGACCTGACC 495
Qy      421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 480
Db      496 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCTCAGTGCATGTAGAGATCCAAAGGGGT 555
Qy      481 AAAAACTACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 540
Db      556 AAAAACTACAGGGGGGAAAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATATGTGC 615
Qy      541 ACCTGGACATGCACTGTCTTGGCAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 600
Db      616 ACCTGGACATGCACTGTCTTGGCAACCAAGAAAGGTGAGTTCAAAATAGACATGTG 675
Qy      601 GTGCTAGCTTTCAGAAAGGCTCCGG 626
Db      676 GTGCTAGCTTTCAGAAAGGCTCCAG 701

```

```

RESULT 6
HUMATCT4      1742 bp      mRNA      linear      PRI 27-APR-1993
LOCUS      HUMATCT4
DEFINITION      Human T-cell surface glycoprotein T4 mRNA, complete cde.
ACCESSION      M12807
VERSION      M12807.1 GI:179141
KEYWORDS      glycoprotein; immunoglobulin super gene family; recognition
antigen; surface antigen.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1
AUTHORS      Maddon,P.J., Liltman,D.R., Godfrey,M., Maddon,D.E., Chess,L. and
Axel,R.
TITLES      The isolation and nucleotide sequence of a cDNA encoding the T cell1
surface protein T4; a new member of the immunoglobulin gene family
JOURNAL      Cell 42 (1), 93-104 (1985)
MEDLINE      85254948
PUBMED      2990730
REFERENCE      2 (bases 153 to 153)
AUTHORS      Liltman,D.R.
JOURNAL      Unpublished (1986)
COMMENT      Original source text: Human peripheral T lymphocyte, cDNA to mRNA,
clone p14B.
T4 is a member of the immunoglobulin supergene family. The T4 mRNA
encodes contiguous V- and J-like elements without the requirement
for DNA recombination events. The V-like domain can be found at
positions 145 to 426 and the J-like domain from 427 to 471. The T4
protein may serve as the specific surface receptor for the AIDS
virus [1].
The revision of the nucleotide at position 153, the amino acid
encoded by 151-153 becomes lysine. This is now regarded to be the
first residue in the mature protein [2].
Location/Qualifiers
1..1742
/organism="Homo sapiens"

```

```

FEATURES
source
1..1742
/organism="Homo sapiens"

```

```

CDS
/mol_type="mRNA"
/db_xref="taxon:9606"
/locus_tag="CD4"
/feature_id="1"
/protein_id="AAA35572.1"
/db_xref="GI:179142"
/translation="MNRGVPFRHLIVLQALPAPATGKRVLAGKGDVETCTAS
OKSIQPHWKNNOIKLGNOSFLTKGSKLNDPDRSLMDGNPLIINKLIE
DSDTICEVDEKREYQVLVGLTANSDHLLQGSLLTLTLPSSPSVOCRSRG
KNIQGGKTLVSQLELDSDGTWCTVLQNKVEFKIDIVLAFKASSIVYKKEBO
VFSFPLAFVETKLTGSGELMWQAEKASSKMTIFDLKNEVSVKWTQDPKQMGK
KLPHLTLPQALPOYAGSGLNLTALAEATGKHQEVNLVVRATOLQNLCEVWGP
SPKMLSLKLNKAKVSKREKAVWLNPEAGMOCILSDSGOVLSENIKVLPTWST
PVQPMALIVLGVAGLLFLFGLGIFPCVCRHRRRQARMQIRLLSEKKTCCPFR
FQKTCSP"
sig_peptide
76..150
/note="T4 surface glycoprotein signal peptide"
mat_peptide
151..1449
/product="T4 surface glycoprotein mature peptide"
ORIGIN
198 bp upstream of RsaI site.
Query Match 43.4%; Score 624.4; DB 9; Length 1742;
Best Local Similarity 99.8%; Pred. No. 1.7e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 135
QY 61 GCAGCACTCAGGAGAAAGAGTGTGTGCGCAAAAAGGGGATACAGTGAAGTACC 120
DB 136 GCAGCACTCAGGAGAAAGAGTGTGTGCGCAAAAAGGGGATACAGTGAAGTACC 195
QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCATCTGAGAAAAACTCCACCAAGATAAG 180
DB 196 TGTACAGCTTCCAGAGAGAGCATACATTCATCTGAGAAAAACTCCACCAAGATAAG 255
QY 181 ATTCGGGAAATCAGGGCTCTTCTTAATTAAGTCCATCCAGCTGAATGATGCGCT 240
DB 256 ATTCGGGAAATCAGGGCTCTTCTTAATTAAGTCCATCCAGCTGAATGATGCGCT 315
QY 241 GACTCAGAGAGAGCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAGAGAGAGCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATAGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACAGAGAGAGTGAATTG 360
DB 376 ATAGAAGACTCAGATATCTTACATCTGTGAAGTGAAGACAGAGAGAGTGAATTG 435
QY 361 CTAGTGTTCGATTTGATCTGCACTGTGACACCTGCTTCAAGGGGAGAGCTGACC 420
DB 436 CTAGTGTTCGATTTGATCTGCACTGTGACACCTGCTTCAAGGGGAGAGCTGACC 495
QY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTGGCCCTCAGTGCATGTAGAGTCCAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTGTAGTGGCCCTCAGTGCATGTAGAGTCCAGGGGT 555
QY 481 AAAAAATACAGAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540
DB 556 AAAAAATACAGAGGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 615
QY 541 ACCTGGAATGACATCTGTCTTGAAGAACAGAGAGTGAATTCAAAATAGCATCTGTG 600
DB 616 ACCTGGAATGACATCTGTCTTGAAGAACAGAGAGTGAATTCAAAATAGCATCTGTG 675
QY 601 GTGTAGCTTTCCAGAGGGCTCCGG 626
DB 676 GTGTAGCTTTCCAGAGGGCTCCAG 701

```

```

LOCUS HUMACT4A 1742 bp DNA linear PRI 24-JUN-1994
DEFINITION Human T4 surface glycoprotein CD4 gene, complete cds.
ACCESSION M35160.1
VERSION M35160.1 GI:179143
KEYWORDS glycoprotein; immunoglobulin super gene family; recognition antigen; surface antigen.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE
1 (bases 1 to 1742)
Hodge, T.W., Sasse, D.R. and McDougal, J.S.
Humans with OKT4-epitope deficiency have a single nucleotide base
change in the CD4 gene, resulting in substitution of TRP240 for
ARG240
JOURNAL Hum. Immunol. 30 (2), 99-104 (1991)
MEDLINE 91216786
PUBMED 1708753
REFERENCE 2 (bases 1 to 1742)
Hodge, T.W.
Direct Submission
Submitted (12-JUN-1990) T.W. Hodge, A-25 Bldg. 1, Rm. 1226, Centers
for Disease Control, 1600 Clifton Rd. Atlanta, GA 30333 USA
COMMENT
A cytosine to thymidine transition at nucleotide position 868
results in substitution of TRP-240 for ARG-240.
FEATURES
source
1..1742
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
/haplotype="OKT4-negative"
76..1452
/gene="CD4"
76..1452
/note="T4 surface glycoprotein precursor"
CDS
76..1452
/gene="CD4"
76..1452
/codon_start=1
/protein_id="AA16069.1"
/db_xref="GI:179144"
/translation="MNRGVPFRHLIVLQALPAPATGKRVLAGKGDVETCTAS
OKSIQPHWKNNOIKLGNOSFLTKGSKLNDPDRSLMDGNPLIINKLIE
DSDTICEVDEKREYQVLVGLTANSDHLLQGSLLTLTLPSSPSVOCRSRG
KNIQGGKTLVSQLELDSDGTWCTVLQNKVEFKIDIVLAFKASSIVYKKEBO
VFSFPLAFVETKLTGSGELMWQAEKASSKMTIFDLKNEVSVKWTQDPKQMGK
KLPHLTLPQALPOYAGSGLNLTALAEATGKHQEVNLVVRATOLQNLCEVWGP
SPKMLSLKLNKAKVSKREKAVWLNPEAGMOCILSDSGOVLSENIKVLPTWST
PVQPMALIVLGVAGLLFLFGLGIFPCVCRHRRRQARMQIRLLSEKKTCCPFR
FQKTCSP"
sig_peptide
76..150
/gene="CD4"
/note="T4 surface glycoprotein signal peptide"
mat_peptide
151..1449
/gene="CD4"
/product="T4 surface glycoprotein mature peptide"
792
/gene="CD4"
/note="g in OKT4 epitope deficiency; c in wt"
ORIGIN
Query Match 43.4%; Score 624.4; DB 9; Length 1742;
Best Local Similarity 99.8%; Pred. No. 1.7e-134;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTGTGCTGCAACTGGCGCTCCCA 135
QY 61 GCAGCACTCAGGAGAAAGAGTGTGTGCGCAAAAAGGGGATACAGTGAAGTACC 120
DB 136 GCAGCACTCAGGAGAAAGAGTGTGTGCGCAAAAAGGGGATACAGTGAAGTACC 195
QY 121 TGTACAGCTTCCAGAGAGAGCATACATTCATCTGAGAAAAACTCCACCAAGATAAG 180

```

|||||  
Db 196 TGTACAGCTCCCGAAGAGATCAATTCACGTGAAAAAACTCCACAGATAAAG 255  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 240  
Db 256 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 315  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300  
Db 316 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 375  
Qy 301 ATAGAACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 360  
Db 376 ATAGAACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 435  
Qy 361 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGAGAGCTACC 420  
Db 436 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGAGAGCTACC 495  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGCATGTAGAGATCCAAAGGGT 480  
Db 496 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGCATGTAGAGATCCAAAGGGT 555  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGGCTCCAGATAGTGC 540  
Db 556 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGGCTCCAGATAGTGC 615  
Qy 541 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATGACATGTG 600  
Db 616 ACCTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATGACATGTG 675  
Qy 601 GTGCTAGCTTCCAGAAAGCCTCCGG 626  
Db 676 GTGCTAGCTTCCAGAAAGCCTCCAG 701

RESULT 8  
LOCUS 108116 1910 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 3 from Patent EP 0325262.  
ACCESSION 108116  
VERSION 108116.1 GI:589172  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 1910)  
AUTHORS Brian,S.D.  
TITLE Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 3 26-JUL-1989;  
FEATURES  
source  
1..1910  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 1910;  
Best Local Similarity 99.8%; Pred.No.1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGCAACTGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGCAACTGCGCTCTCCCA 170  
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 230  
Qy 121 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAACTCCACAGATAAAG 180  
Db 231 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAACTCCACAGATAAAG 290  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 240

|||||  
Db 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 350  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300  
Db 351 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 410  
Qy 301 ATAGAACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 360  
Db 411 ATAGAACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGAATTG 470  
Qy 361 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGAGAGCTACC 420  
Db 471 CTAGTGTTCGATTTGACTGCAACTGTGACACCCACTGCTTCAAGGGGAGAGCTACC 530  
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGCATGTAGAGATCCAAAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGTAGTACCCCTCAGTGCATGTAGAGATCCAAAGGGT 590  
Qy 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGGCTCCAGATAGTGC 540  
Db 591 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTGAGCTGAGGCTCCAGATAGTGC 650  
Qy 541 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGATTCAAAATGACATGTG 600  
Db 651 ACCTGACATGCACTGTCTTGCAGAAACAGAAAGGTGAGATTCAAAATGACATGTG 710  
Qy 601 GTGCTAGCTTCCAGAAAGCCTCCGG 626  
Db 711 GTGCTAGCTTCCAGAAAGCCTCCAG 736

RESULT 9  
LOCUS 108115 3133 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0325262.  
ACCESSION 108115  
VERSION 108115.1 GI:589171  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 3133)  
AUTHORS Brian,S.D.  
TITLE Cloned genes encoding IG-CD4 fusion proteins and the use thereof  
JOURNAL Patent: EP 0325262-A2 1 26-JUL-1989;  
FEATURES  
source  
1..3133  
/organism="unknown"  
/mol\_type="unassigned DNA"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 6; Length 3133;  
Best Local Similarity 99.8%; Pred.No.1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGCAACTGCGCTCTCCCA 60  
Db 111 ATGAACCGGGAGTCCCTTTTAGCACTTGCTTGCTGTGCTGCAACTGCGCTCTCCCA 170  
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120  
Db 171 GCAGCCACTCAGGGAAGAAAGTGTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 230  
Qy 121 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAACTCCACAGATAAAG 180  
Db 231 TGTACAGCTTCCAGAAAGAGATCAATTCACCTGGAAGAACTCCACAGATAAAG 290  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 240  
Db 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCAATCCAGCTGAATGATCGGCT 350  
Qy 241 GACTCAAGAAAGCCTTTGGGACCAAGAACTTCCCGCTGATCATCAAGATCTTAAG 300

Db 351 GACTCAAGAAAGACCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
Qy 301 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
Db 411 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 470  
Qy 361 CTAGTGTGGATTGATCTGCACTGTGACACCACTGCTTCAAGGGGCAAGACCTGACC 420  
Db 471 CTAGTGTGGATTGATCTGCACTGTGACACCACTGCTTCAAGGGGCAAGACCTGACC 530  
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGT 480  
Db 531 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGT 590  
Qy 481 AAAAATATCAGAGGGGGGAAAGACCTCTCTCTGCTCTCACTGAGAGTCCAGATAGTGC 540  
Db 591 AAAAATATCAGAGGGGGGAAAGACCTCTCTCTGCTCTCACTGAGAGTCCAGATAGTGC 650  
Qy 541 ACCTGAGATGACCTGCTTGTGAGACCAAGAGAGTGAAGTCAATAGCATCTG 600  
Db 651 ACCTGAGATGACCTGCTTGTGAGACCAAGAGAGTGAAGTCAATAGCATCTG 710  
Qy 601 GTGCTAGCTTTCCAGAGGCTCTCCG 626  
Db 711 GTGCTAGCTTTCCAGAGGCTCTCCAG 736

RESULT 10  
AY438650 7533 bp DNA linear SYN 11-NOV-2003  
LOCUS Cloning vector pMACS1Bac, complete sequence.  
DEFINITION AY438650  
ACCESSION AY438650  
VERSION AY438650.1 GI:38197734  
KEYWORDS  
SOURCE Cloning vector pMACS1Bac  
ORGANISM Cloning vector pMACS1Bac  
artificial sequences; vectors.  
REFERENCE 1 (bases 1 to 7533)  
AUTHORS Phillips,B., Forstner,M. and Mayr,L.M.  
TITLE A baculovirus expression system for magnetic sorting of infected cells and enhanced titer determination  
JOURNAL Unpublished  
REFERENCE 2 (bases 1 to 7533)  
AUTHORS Phillips,B., Forstner,M. and Mayr,L.M.  
TITLE Direct Subdivision  
JOURNAL Submitted (16-OCT-2003) IDC / BSO1, Novartis Pharma AG, WSJ-88.601, Basel CH-4002, Switzerland  
FEATURES  
SOURCE Location/Qualifiers  
1.7533  
/organism="Cloning vector pMACS1Bac"  
/mol\_type="other DNA"  
/db\_xref="taxon:254251"  
689..1549  
/gene="bia"  
689..1549  
/gene="bia"  
/codon\_start=1  
/transl\_table=1  
/product="beta-lactamase"  
/protein\_id="AAR13800.1"  
/db\_xref="GI:38197735"  
/translation="MSIQHFRVALIPFPAFLPVFAPHELVKVKADQEDQGARVGY  
LELDNSGLIESFRPEERFPWMTSTFKVLLCGAVSRIDAGQELGRIRHVSQNDLVE  
YSPVTEKHLIDMTRELCSAAITMSDVTALLLTTIGSPKELTAPFLHNMGDHTRL  
DRMEPLNBAIPNDERDITMPVAMAATTARKLLTGLLTLASQOOLIDMWEADKVA  
GLPLLRSLAPAGFWIADKSGAGBGRSGRIIALLGPGRSRIIVITTSQATMDERNQIA  
ELGASLTKHW"  
complement (2902..3435)  
/gene="aacC1"  
complement (2902..3435)  
/gene="aacC1"  
/codon\_start=1

/transl\_table=1  
/product="gentamicin-3-acetyl transferase"  
/protein\_id="AAR13801.1"  
/db\_xref="GI:38197736"  
/translation="MLRSSNDVTQOGRPKTKLGSSMGIRFTRCHLPDQVSKMBAL  
DLFGEFGVATYSOHOPDSIDYGLNLSKPTLIALAPQEAVALAAVYPRFPO  
RSEIYIDLAVSGEHRROGIALNLKHEMNLGAVIYVQADYGDPAVALTKL  
GIREVHFHDIDESTAT"  
complement (4301..6394)  
/gene="CD4-EGFP"  
complement (4301..6394)  
/note="CD4 and enhanced GFP fusion protein"  
/codon\_start=1  
/transl\_table=1  
/product="CD4-EGFP"  
/protein\_id="AAR13802.1"  
/db\_xref="GI:38197737"  
/translation="MNRGVPPRHLLVVLQALLPAATQKRVVLGKKGDTVELCTAS  
OKKSIQPHKNSNQIKITLNGSGSLITKPSKLNDPADRSLSMDGAPFLITKNIKIE  
DSDIYICEVDQKEVQLVFGLTANSIDHLOGLTLTLESPPGSSPSVOCSPRG  
KNIQGGKTLVSQLELODGTWCTVLQNKVEPKIDIVLAPOKASIVYKKEGO  
VEFSPLAFVTEKLTSGSELWMOARASSSKSMITFDLKNVESYKRVYODPKLQMGK  
KLPLHLTLPOLPOYAGSGNLTALAEATGKLEHVNLVNRAATOLKULCEWNGPT  
SPKMLSLKLENKKAIVSRKRAVWVLPKAGMOCILSDSGOVLNENIKVLPWST  
PVQPMALIVLGSVAGLLPLIGLGIFFCYCRHRRQARMSQIKLLSEKTKCQCPHR  
FOKTSPIVSKEBELFTGVVPIVELDQVNGHRSVSGEBGATGKLTFTICT  
TGKLPVWPVTLVTLTVGVQCFRYPDMKQHPDFKSMPEGVVERTIFPDNDNYK  
TRAEVKFEGLTVNRIELKGIIDFKEDGNIIGKLEYNLSHVYIMADOKNGIKYFN  
KIRHNIEDGSVOLADHYQONTPIGDGPVILPDPNHYLSQALSQDPNEKRDMVLEF  
VTAAGITLGMELVYK"

## ORIGIN

Query Match 43.4%; Score 624.4; DB 12; Length 7533;  
Best Local Similarity 99.8%; Pred. No. 1.7e-134;  
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 60  
Db 6394 ATGAACCGGGAGTCCCTTTTAGGCACTTCTGTGTGCTGCAACTGGCGCTCTCCCA 6335  
Qy 61 GCAGCCTCAGGGGAAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTCAACC 120  
Db 6334 GCAGCCTCAGGGGAAAGAGTGTGCTGGGCAAAAAGGGGATACAGTGAAGTCAACC 6275  
Qy 121 TGTACAGCTTCCCAAGAGAGCATACATTCACCTGGAATACTCCAAACAGATTAAG 180  
Db 6274 TGTACAGCTTCCCAAGAGAGCATACATTCACCTGGAATACTCCAAACAGATTAAG 6215  
Qy 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATATGCGCT 240  
Db 6214 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATATGCGCT 6155  
Qy 241 GACTCAAGAGAGCTTTGGGACCAAGAACTCCCTCGATATCAAGAACTTTAAG 300  
Db 6154 GACTCAAGAGAGCTTTGGGACCAAGAACTCCCTCGATATCAAGAACTTTAAG 6095  
Qy 301 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 360  
Db 6094 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTG 6035  
Qy 361 CTAGTGTGGATTGATCTGCACTGTGACACCACTGCTTCAAGGGGCAAGACCTGACC 420  
Db 6034 CTAGTGTGGATTGATCTGCACTGTGACACCACTGCTTCAAGGGGCAAGACCTGACC 5975  
Qy 421 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGT 480  
Db 5974 CTGACCTTGGAGAGCCCCCTGGTAGTACCCCTCAGTCAATGTAGAGTCCAGAGGGT 5915  
Qy 481 AAAAATATCAGAGGGGGGAAAGACCTCTCTCTGCTCTCACTGAGAGTCCAGATAGTGC 540  
Db 5914 AAAAATATCAGAGGGGGGAAAGACCTCTCTCTGCTCTCACTGAGAGTCCAGATAGTGC 5855

QY 541 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
DB 5854 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 5795  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 5794 GTGCTAGCTTTCCAGAAAGGCTCCAG 5769

RESULT 11  
LOCUS 106223 1273 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0313377.  
ACCESSION 106223  
VERSION 106223.1 GI:590426  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1273)  
AUTHORS Deen,K.C., Folena-Wasserman,G.M., Inacker,R.H. and Sweet,R.W.  
TITLE Process for purification of soluble T4  
JOURNAL Patent: EP 0313377-A1 1 26-APR-1989;  
FEATURES  
source location/Qualifiers  
1..1273  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 43.2%; Score 622.8; DB 6; Length 1273;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGAGTGAAGTGGGCTCTCCCA 60  
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGAGTGAAGTGGGCTCTCCCA 135  
QY 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGATAGAGTGAATGACC 120  
DB 136 GCAGCCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGATAGAGTGAATGACC 195  
QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCATGGAATACTCCACAGATTAAG 180  
DB 196 TGTACAGCTTCCAGAAAGAGCATTCATTCCATGGAATACTCCACAGATTAAG 255  
QY 181 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAAGTCCAGAGTGAATGAGGCT 240  
DB 256 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAAGTCCAGAGTGAATGAGGCT 315  
QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 316 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
QY 301 ATGAAAGACTCAGATTAATCTTATCATCTGGAAGTGAAGCAAGAGAGAGTGCATTG 360  
DB 376 ATGAAAGACTCAGATTAATCTTATCATCTGGAAGTGAAGCAAGAGAGAGTGCATTG 435  
QY 361 CTAGTGTTCGATTTGACTGCACTGCAACCTGATCCAGGGGCAAGGCTGACC 420  
DB 436 CTAGTGTTCGATTTGACTGCACTGCAACCTGATCCAGGGGCAAGGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 480  
DB 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 555  
QY 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTCAGTGAAGTCCAGATAGTGGC 540  
DB 556 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTCAGTGAAGTCCAGATAGTGGC 615  
QY 541 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
DB 616 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 675

QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 12  
LOCUS 107147 1273 bp DNA linear PAT 02-DEC-1994  
DEFINITION Sequence 1 from Patent EP 0313356.  
ACCESSION 107147  
VERSION 107147.1 GI:590041  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1273)  
AUTHORS Arthos,J., Clark,P.E., Formwald,J.A., Brawer,M.E., Deen,K.C.,  
Gorman,J.A., Sathie,G.M., Sweet,R.W. and Taylor,D.P.  
TITLE Expression of HIV binding proteins  
JOURNAL Patent: EP 0313356-A2 1 06-SEP-1989;  
FEATURES  
source location/Qualifiers  
1..1273  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Query Match 43.2%; Score 622.8; DB 6; Length 1273;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGAGTGAAGTGGGCTCTCCCA 60  
DB 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGCTGCTGAGTGAAGTGGGCTCTCCCA 135  
QY 61 GCAGCCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGATAGAGTGAATGACC 120  
DB 136 GCAGCCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGATAGAGTGAATGACC 195  
QY 121 TGTACAGCTTCCAGAAAGAGCATTCATTCCATGGAATACTCCACAGATTAAG 180  
DB 196 TGTACAGCTTCCAGAAAGAGCATTCATTCCATGGAATACTCCACAGATTAAG 255  
QY 181 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAAGTCCAGAGTGAATGAGGCT 240  
DB 256 ATTCTGGGAATCAGGGCTCCTTCTTAAGTCAAGTCCAGAGTGAATGAGGCT 315  
QY 241 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 316 GACTCAAGAAAGCCCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 375  
QY 301 ATGAAAGACTCAGATTAATCTTATCATCTGGAAGTGAAGCAAGAGAGAGTGCATTG 360  
DB 376 ATGAAAGACTCAGATTAATCTTATCATCTGGAAGTGAAGCAAGAGAGAGTGCATTG 435  
QY 361 CTAGTGTTCGATTTGACTGCACTGCAACCTGATCCAGGGGCAAGGCTGACC 420  
DB 436 CTAGTGTTCGATTTGACTGCACTGCAACCTGATCCAGGGGCAAGGCTGACC 495  
QY 421 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 480  
DB 496 CTGACCTTGGAGAGCCCTCTGTAGTAGCCCTCAGTGCATGTAGAGTCCAAAGGGT 555  
QY 481 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTCAGTGAAGTCCAGATAGTGGC 540  
DB 556 AAAAACAATACAGGGGGGAGAACCTCTCCGTGTCTCAGTGAAGTCCAGATAGTGGC 615  
QY 541 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 600  
DB 616 ACCTGACATGACTGTCTTGACAGAACAGAAAGGTGAGATTCAAAATATGACATCGTG 675  
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626  
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

RESULT 13  
LOCUS AR067943 1304 bp DNA linear PAT 29-SEP-1999  
DEFINITION Sequence 28 from patent US 5851828.  
ACCESSION AR067943  
VERSION AR067943.1 GI:5999165  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1304)  
AUTHORS Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
receptor-bearing cells  
JOURNAL Patent: US 5851828-A 28 22-DEC-1998;  
FEATURES Location/Qualifiers  
source 1..1304  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 170  
QY 61 GCAGCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAGGGAGATGAGTGAACCTGACC 120  
DB 171 GCAGCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAGGGAGATGAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 180  
DB 231 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
QY 211 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
DB 221 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATGATGCGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATGATGCGCT 350  
QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 351 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
QY 301 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
DB 411 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
QY 411 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
DB 421 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
QY 361 CTAGTGTTCGATTTGATGCTGCAACTCTGACCACTGCTTCAAGGGGCAAGCTGACC 420  
DB 471 CTAGTGTTCGATTTGATGCTGCAACTCTGACCACTGCTTCAAGGGGCAAGCTGACC 530  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 480  
DB 531 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 590  
QY 481 AAAAATATACAGGGGGGAGAGACCTCTCCGCTGCTCAGCTGAGACTCAGAGATGAGGC 540  
DB 591 AAAAATATACAGGGGGGAGAGACCTCTCCGCTGCTCAGCTGAGACTCAGAGATGAGGC 650  
QY 541 ACCTGAGATGACATCTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATCTG 600  
DB 651 ACCTGAGATGACATCTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATCTG 710  
QY 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626  
DB 711 GTGCTAGCTTCCAGAAAGGCTCCAG 736

RESULT 14

AR166802  
LOCUS AR166802 1304 bp DNA linear PAT 17-OCT-2001  
DEFINITION Sequence 28 from patent US 6284240.  
ACCESSION AR166802  
VERSION AR166802.1 GI:16243143  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1304)  
AUTHORS Seed, B., Banapour, B., Romeo, C. and Kolanus, W.  
TITLE Targeted cytolysis of HIV-infected cells by chimeric CD4  
receptor-bearing cells  
JOURNAL Patent: US 6284240-A 28 04-SEP-2001;  
FEATURES Location/Qualifiers  
source 1..1304  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 43.2%; Score 622.8; DB 6; Length 1304;  
Best Local Similarity 99.7%; Pred. No. 4.1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 60  
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTCTGTGCTGCAACTGGCGCTCTCCCA 170  
QY 61 GCAGCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAGGGAGATGAGTGAACCTGACC 120  
DB 171 GCAGCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAGGGAGATGAGTGAACCTGACC 230  
QY 121 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 180  
DB 231 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
QY 211 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
DB 221 TGTACAGCTTCCCAAGAGAGCATACATTCCTGGAATACTCCAACTGATAAAG 290  
QY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATGATGCGCT 240  
DB 291 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATGATGCGCT 350  
QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
DB 351 GACTCAAGAGAGGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 410  
QY 301 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 360  
DB 411 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
QY 411 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
DB 421 ATAGAAGACTCAGATTAATTAATCTGTGAAGTGAAGACCAAGAGAGGATGCAATTG 470  
QY 361 CTAGTGTTCGATTTGATGCTGCAACTCTGACCACTGCTTCAAGGGGCAAGCTGACC 420  
DB 471 CTAGTGTTCGATTTGATGCTGCAACTCTGACCACTGCTTCAAGGGGCAAGCTGACC 530  
QY 421 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 480  
DB 531 CTGACCTTGGAGAGGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGGCT 590  
QY 481 AAAAATATACAGGGGGGAGAGACCTCTCCGCTGCTCAGCTGAGACTCAGAGATGAGGC 540  
DB 591 AAAAATATACAGGGGGGAGAGACCTCTCCGCTGCTCAGCTGAGACTCAGAGATGAGGC 650  
QY 541 ACCTGAGATGACATCTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATCTG 600  
DB 651 ACCTGAGATGACATCTGTCTTGCAGAACCAAGAAAGGTGAGATTCAAAATAGACATCTG 710  
QY 601 GTGCTAGCTTCCAGAAAGGCTCCGG 626  
DB 711 GTGCTAGCTTCCAGAAAGGCTCCAG 736

RESULT 15

AR062468 1389 bp DNA linear PAT 29-SEP-1999  
LOCUS AR062468  
DEFINITION Sequence 2 from patent US 5843728.

ACCESSION AR062468  
VERSION AR062468.1 GI:5990159  
KEYWORDS  
SOURCE  
ORGANISM Unknown.  
REFERENCE Unclassified.  
1 (bases 1 to 1389)  
AUTHORS Seed, B., Romeo, C. and Kolanus, W.  
TITLE Redirection of cellular immunity by receptor chimeras  
JOURNAL Patent: US 5843728-A 2 01-DEC-1998;  
FEATURES Location/Qualifiers  
1..1389  
source

ORIGIN  
/organism="unknown"  
/mol\_type="unassigned DNA"

Query Match 43.2%; Score 622.8; DB 6; Length 1389;  
Best Local Similarity 99.7%; Pred. No. 4,1e-134;  
Matches 624; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGCTGGTGCACACTGGGGCTCTCCCA 60  
Db 1 ATGAACCGGGAGTCCCTTTTAAAGCACTTGCTGGTGCACACTGGGGCTCTCCCA 60  
QY 61 GCAGCCACTCAGGGAAAGAAAGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
Db 61 GCAGCCACTCAGGGAAACAAAGTGTGTCTGGGCAAAAAAGGGATACAGTGAACCTGACC 120  
QY 121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAAAAAATCCAAACGATTAAG 180  
Db 121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAAAAAATCCAAACGATTAAG 180  
QY 181 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATGCGGCT 240  
Db 181 ATTCTGGGAATCAGGGCTCTTCTTAAGTAAAGTCCATCCAGCTGAATGCGGCT 240  
QY 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
Db 241 GACTCAAGAAAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300  
QY 301 ATAGAGACTCAGATTAATCTGTTGAAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
Db 301 ATAGAGACTCAGATTAATCTGTTGAAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
QY 361 CTAGTGTGGATGATGCGCACTGACACCCACCGTTCAGGGGCAAGCCTGACC 420  
Db 361 CTAGTGTGGATGATGCGCACTGACACCCACCGTTCAGGGGCAAGCCTGACC 420  
QY 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATGTAGAGTCCAAGGGGT 480  
Db 421 CTGACCTTGGAGAGCCCCCTGTAGTAGCCCCCTCAGTGCATGTAGAGTCCAAGGGGT 480  
QY 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCACTGAGCTCCAGATATGCGC 540  
Db 481 AAAAACAATACAGGGGGGAAAGACCTCTCCGTGTCTCACTGAGCTCCAGATATGCGC 540  
QY 541 ACCCTGACATGCACTGCTTGGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
Db 541 ACCCTGACATGCACTGCTTGGCAAGACCAAGAAAGGTGAGTCAAAATAGACATGCTG 600  
QY 601 GTGCTAGCTTCCAGAGGCTCCGG 626  
Db 601 GTGCTAGCTTCCAGAGGCTCCAG 626

Search completed: November 22, 2004, 12:51:24  
Job time : 6284 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 21, 2004, 14:00:14 ; Search time 754 Seconds  
(without alignments)  
10025.419 Million cell updates/sec

Title: US-09-936-702-4

Perfect score: 1440

Sequence: 1 atgacgcggggagcgcctt.....gtcagcagctacataactg 1440

Scoring table: IDENTITY\_NUC  
Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Listing first 45 summaries

Database : N\_Geneseq\_23Sep04:\*

1: geneeqn1980s:\*  
2: geneeqn1990s:\*  
3: geneeqn2000s:\*  
4: geneeqn2001as:\*  
5: geneeqn2001bs:\*  
6: geneeqn2002as:\*  
7: geneeqn2002bs:\*  
8: geneeqn2003as:\*  
9: geneeqn2003bs:\*  
10: geneeqn2003cs:\*  
11: geneeqn2003ds:\*  
12: geneeqn2004s:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1440	100.0	1440	3	Aa54045 sCD4-SCFV
2	625	43.4	8911	12	ADN07735 Expressio
3	624.4	43.4	1377	4	AAf82582 Human CD4
4	624.4	43.4	1415	2	AAQ38761 sCD4-L2.1
5	624.4	43.4	1419	10	ADA44806 CD4/TCR C
6	624.4	43.4	1421	2	AAQ38759 sCD4-L1.1
7	624.4	43.4	1448	2	AAQ38759 sCD4-HAP
8	624.4	43.4	1714	3	AA550662 DNA encod
9	624.4	43.4	1714	3	AA244063 Human fus
10	624.4	43.4	1714	3	AA248203 DNA seque
11	624.4	43.4	1742	3	AA35205 Human ade
12	624.4	43.4	1742	3	AAf21327 Human low
13	624.4	43.4	1742	4	AA165462 Nucleotid
14	624.4	43.4	1742	10	AB297021 Human nuc
15	624.4	43.4	1742	11	AD131687 Human nuc
16	624.4	43.4	1742	11	ABD20870 Human pul
17	624.4	43.4	1910	3	AA550663 DNA encod
18	624.4	43.4	1910	3	AA244064 Human fus
19	624.4	43.4	1910	3	AA248204 DNA seque
20	624.4	43.4	2150	3	AA248205 DNA seque
21	624.4	43.4	2589	1	AA90357 Genetic c

22	624.4	43.4	2589	3	AA550661 DNA encod
23	624.4	43.4	2589	3	AA244062 Human fus
24	624.4	43.4	2589	3	AA248202 DNA seque
25	624.4	43.4	3133	1	AA90356 Genetic c
26	624.4	43.4	3133	3	AA550660 DNA encod
27	624.4	43.4	3133	3	AA244061 Human fus
28	624.4	43.4	3133	3	AA248201 DNA seque
29	624.4	43.4	6019	3	AA235208 Human ade
30	624.4	43.4	6019	3	AAf21330 Human low
31	624.4	43.4	6019	10	AB297024 Human nuc
32	624.4	43.4	6019	11	ABD20873 Human pul
33	622.8	43.2	1304	2	AAQ96103 CD4 doma
34	622.8	43.2	1304	2	AAf10797 CD4 D1-D4
35	622.8	43.2	1389	2	AAQ28705 DNA encod
36	622.8	43.2	1389	2	AAQ96123 T-cell re
37	622.8	43.2	1389	2	AAf10802 CD4:gamma
38	622.8	43.2	1389	2	AAf36759 CD4:FC re
39	622.8	43.2	1389	2	AAV70158 Chimeric
40	622.8	43.2	1599	2	AAQ28706 DNA encod
41	622.8	43.2	1599	2	AAf10803 CD4:eta f
42	622.8	43.2	1599	2	AAf36760 CD4:T-cel
43	622.8	43.2	1599	2	AAV70157 Chimeric
44	622.8	43.2	1727	2	AAQ96124 T-cell re
45	622.8	43.2	1728	2	AAQ28704 DNA encod

#### ALIGNMENTS

RESULT 1	AAA54045	standard: DNA; 1440 BP.
XX	AAA54045	
AC	AAA54045	
XX		
DT	08-FEB-2001	(first entry)
XX		
DE	sCD4-SCFV(17b)	HIV single chain antibody fusion protein.
XX		
KW	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;	
KW	acquired immune deficiency syndrome; neutralisation; infection;	
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;	
KW	binding domain; single chain antibody; chimera; chimeric protein; ds.	
OS	Human immunodeficiency virus.	
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	CDS	1..1440
FT		/*tag= a
FT		/product= "CD4-SCFV(17b) fusion protein."
FT		/note= "This sequence does not fully encode the cross
FT		referenced protein given in GENESQ record AAB00158. That
FT		protein sequence also comprises a 28 amino acid C-
FT		terminal peptide"
XX		
PN	WO20055207-A1.	
XX		
PD	21-SEP-2000.	
XX		
PF	16-MAR-2000; 2000WO-US006946.	
XX		
PR	16-MAR-1999; 99US-0124681P.	
XX		
PA	(USSH ) US NAT INST OF HEALTH.	
XX		
PI	Berger EA, Del Castillo CM;	
XX		
DR	WPI, 2000-638183/61.	
DR	P-PSDB; AAB00158.	
XX		
PT	Novel neutralizing bispecific fusion proteins effective in viral such as	
PT	HIV neutralization, comprises two different binding domains, inducing-	

binding domain and induced-binding domain functionally linked by linker.

Claim 36; Page 47-48; 55pp; English.

CD4-8CFv(17b) is a neutralising bispecific fusion protein capable of binding to two sites of its target protein. The protein comprises a first binding domain capable of binding to an inducing site on the target protein, a second binding domain capable of forming neutralising complex with an induced epitope of the target protein and a linker connecting the binding domains. CD4-8CFv(17b) comprises a soluble CD4 fragment (containing domains D1 and D2) fused to a single chain Fv portion of antibody 17b via a linker. CD4-8CFv(17b), its variant, analogue or mimetic is used for inactivating gp120 protein of HIV, and for neutralising HIV. It is also used for blocking and preventing the binding of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte CD4 and for inhibiting HIV replication. The chimeric proteins is therefore useful for treating HIV infection and also AIDS. It is are particularly useful in the prevention of infection during or immediately after HIV exposure (e.g., mother/infant transmission, post-exposure prophylaxis, and as a topical inhibitor) and for providing long term resistance to HIV infections and AIDS. Gene therapy is used to secrete the bispecific protein at mucosal surfaces, such as the vaginal, rectal or oral mucosa. The fusion proteins is highly potent, broadly cross-reactive with neutralising antibody with high in vivo activity and no Fc-mediated undesirable targeting properties. When the fusion protein is substantially derived from human proteins, it has minimal immunogenicity and toxicity in humans which is of great value in prevention of infection during or immediately after HIV exposure

Sequence 1440 BP; 345 A; 354 C; 452 G; 289 T; 0 U; 0 Other;

Query Match 100.0%; Score 1440; DB 3; Length 1440; Best Local Similarity 100.0%; Pred. No. 1.4e-311; Matches 1440; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 ATGAACCGGAGAGTCCCTTTAGGCACTTGTGTGTGTCGCAATGGCGCTCTCCCA 60  
1 ATGAACCGGAGAGTCCCTTTAGGCACTTGTGTGTGTCGCAATGGCGCTCTCCCA 60  
61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTCGCAAAAGGGGATACAGTGAACCTAAC 120  
61 GCAGCCCTCAGGGAGAAAGTGTGTGTGTCGCAAAAGGGGATACAGTGAACCTAAC 120  
121 TGTACAGCTTCCAGAGAGAGCATCAATTCCACTGGAAGAACTCCACCAAGATTAAG 180  
121 TGTACAGCTTCCAGAGAGAGCATCAATTCCACTGGAAGAACTCCACCAAGATTAAG 180  
181 ATTCTGGGAATCAGGCTCTCTTTAACTTAAGGTCCATCAAGCTGAATGATCGGCT 240  
181 ATTCTGGGAATCAGGCTCTCTTTAACTTAAGGTCCATCAAGCTGAATGATCGGCT 240  
241 GACTCAGAGAAAGCTTTGGGACCAAGAAACTCCCGTCGATCATCAAGAACTTAAG 300  
241 GACTCAGAGAAAGCTTTGGGACCAAGAAACTCCCGTCGATCATCAAGAACTTAAG 300  
301 ATAGAAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
301 ATAGAAAGCTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTCAATTG 360  
361 CTAGTGTTCGAGATTGACTGCCAATCTGACACCACTGCTTCAGGGGACAGCTTAAC 420  
361 CTAGTGTTCGAGATTGACTGCCAATCTGACACCACTGCTTCAGGGGACAGCTTAAC 420  
421 CTGACCTTGGAGAGCCCCCTGGTAGAGCCCTCAGTGCATATAGAGAGTCCAAAGG 480  
421 CTGACCTTGGAGAGCCCCCTGGTAGAGCCCTCAGTGCATATAGAGAGTCCAAAGG 480  
481 AAAAACAATACAGGGGGAGAGACCTCTCCGTCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
481 AAAAACAATACAGGGGGAGAGACCTCTCCGTCGTCTCAGCTGAGCTCCAGATAGTGGC 540  
541 ACCTGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600  
541 ACCTGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600

541 ACCTGACATGCACTGTCTTGCAAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600  
601 GTGCTAGCTTTCAGAAAGGCTCCGGAGGTGCGGATAGTGGGGAGCGGTTCCAGCGGA 660  
601 GTGCTAGCTTTCAGAAAGGCTCCGGAGGTGCGGATAGTGGGGAGCGGTTCCAGCGGA 660  
661 GGTGATCCGGTGGCGAGGGGTGGGCGGGGTGGAAGCGGGGTGGCGCTCCGAGGC 720  
661 GGTGATCCGGTGGCGAGGGGTGGGCGGGGTGGAAGCGGGGTGGCGCTCCGAGGC 720  
721 GAGGTTCAAGATGTCAGTGTCTGAGTCTGGGGCTGAGGTGAAGAAAGCTGGGTCTCG 780  
721 GAGGTTCAAGATGTCAGTGTCTGAGTCTGGGGCTGAGGTGAAGAAAGCTGGGTCTCG 780  
781 GTAAGGTCCTCTGCAAGGCTCTGAGAGACCTTCATCAGATATAGTTTACCTGGGTG 840  
781 GTAAGGTCCTCTGCAAGGCTCTGAGAGACCTTCATCAGATATAGTTTACCTGGGTG 840  
841 CGAAGGCCCCCTGACAGAGGCTTTGATGGATGGAAAGATCATCATCTATTGATGA 900  
841 CGAAGGCCCCCTGACAGAGGCTTTGATGGATGGAAAGATCATCATCTATTGATGA 900  
901 GCACATACGCAACCGACCTCCAGGCGAGAGTACAGATTACCGGACAAAGTCCAGAGC 960  
901 GCACATACGCAACCGACCTCCAGGCGAGAGTACAGATTACCGGACAAAGTCCAGAGC 960  
961 ACAATCACTGAGCTGCGGAACTTAAGATCTGACGATCGGCGGTATTTCTGTGGC 1020  
961 ACAATCACTGAGCTGCGGAACTTAAGATCTGACGATCGGCGGTATTTCTGTGGC 1020  
1021 GAGGTGACAGGAGAGAGCGGAGCGAGAGGGGAAATATGATTAATGAGGTTCTGAACAT 1080  
1021 GAGGTGACAGGAGAGAGCGGAGCGAGAGGGGAAATATGATTAATGAGGTTCTGAACAT 1080  
1081 TGGGGCCAGGAAACCTTGTCACAGTCACTCAGTGTGGGTGGCTCCGAGGTGTGGG 1140  
1081 TGGGGCCAGGAAACCTTGTCACAGTCACTCAGTGTGGGTGGCTCCGAGGTGTGGG 1140  
1141 AGCGGTGGCGGAGATCTGAATCTGAGTTGACCAAGTCTCCAGCCACCTGTCTGTGCT 1200  
1141 AGCGGTGGCGGAGATCTGAATCTGAGTTGACCAAGTCTCCAGCCACCTGTCTGTGCT 1200  
1201 CCAAGGAAAGAGCCACCTCTCTGCAAGGCGCAGTGAAGTGTAGTGAAGCTTAAGCC 1260  
1201 CCAAGGAAAGAGCCACCTCTCTGCAAGGCGCAGTGAAGTGTAGTGAAGCTTAAGCC 1260  
1261 TGGTACCAAGAAACCTTGCCAGGCTCCAGGCTCCTCATATATGTGTGATCCACAGG 1320  
1261 TGGTACCAAGAAACCTTGCCAGGCTCCAGGCTCCTCATATATGTGTGATCCACAGG 1320  
1321 GCCACCGGTGTCCACGCAAGGTTCAGTGGCAGTGGGTCTGGGGCAGAAATCATCTCACC 1380  
1321 GCCACCGGTGTCCACGCAAGGTTCAGTGGCAGTGGGTCTGGGGCAGAAATCATCTCACC 1380  
1381 ATCAGCAGCTGCAAGTGTGAAGATTTTGAGTTTATCTGTGACAGATCAATAACTGG 1440  
1381 ATCAGCAGCTGCAAGTGTGAAGATTTTGAGTTTATCTGTGACAGATCAATAACTGG 1440

RESULT 2  
ADN07735  
ID ADN07735 standard; DNA; 8911 BP.  
XX  
AC ADN07735;  
XX  
DT 15-JUL-2004 (first entry)  
XX  
DE Expression vector DNA.  
XX  
KW Immunogenic complex; gp120; CD4; HIV; HIV infection; anti-HIV; circular;  
KM cyclic; ds.  
XX  
OS Synthetic.

```

XX US2004076636-A1.
XX
XX
XX
XX 22-APR-2004.
XX
XX
XX 02-JUL-2003; 2003US-00612192.
XX
XX 07-MAY-1993; 93US-00060926.
XX 06-MAY-1994; 94MO-US005020.
XX 20-DEC-1995; 95US-00464680.
XX 11-MAY-1996; 98US-00075544.
XX 07-JAN-2000; 2000US-00479675.
XX 17-JUL-2001; 2001US-00905962.
XX
XX (PALR/) PAL R.
XX (MARK/) MARKHAM P.
XX (KEEN/) KEEN T.
XX (WHIT/) WHITNEY S.
XX (KALY/) KALYANARAMAN V S.
XX
XX Pal R, Markham P, Keen T, Whitney S, Kalyanaraman VS;
XX
XX WPI; 2004-387924/36.
XX
XX Immunogenic complex, useful for treating HIV infections, comprises gp120
XX covalently bonded to CD4 equivalent molecule, fragment of CD4 or its
XX equivalent.
XX
XX Disclosure; SEQ ID NO 3; 26pp; English.
XX
XX The invention relates to an immunogenic complex comprising gp120
XX covalently bonded to a CD4 equivalent molecule, fragment of CD4 or its
XX equivalent. The invention also relates to a composition comprising the
XX complex and a carrier, an antibody reactive with the complex, an
XX immortalised cell line that produces the complex, a method of detecting
XX the HIV antigen in a test fluid, involving contacting the test fluid with
XX an antibody raised against the immunogenic complex and detecting the
XX presence of immune complexes formed between the antigen in the test fluid
XX and the antibody, and a vaccine comprising an immunogenically effective
XX amount of the immunogenic complex. The immunogenic complex is useful for
XX raising neutralising antibodies against HIV, which involves administering
XX the complex to a subject in a carrier, and for treating HIV infections.
XX This sequence represents expression vector DNA used in the scope of the
XX invention.
XX
XX Sequence 8911 BP; 2188 A; 2221 C; 2227 G; 2275 T; 0 U; 0 Other;
XX
XX Query Match 43.4%; Score 625; DB 12; Length 8911;
XX Best Local Similarity 100.0%; Pred. No. 1.7e-129;
XX Matches 625; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCAATGCGGCTCTCCCA 60
XX Db 3815 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGCAATGCGGCTCTCCCA 3874
XX
XX 61 GCAGCCTCAGGAGGAAAGAACTGCTGGGCAAAAAAGGAGTACGATGAACTGACC 120
XX Db 3875 GCAGCCTCAGGAGGAAAGAACTGCTGGGCAAAAAAGGAGTACGATGAACTGACC 3934
XX
XX 121 TGTACAGCTTCCGAGAGAGATACATTCACCTGGAATACTCCACAGATAAAG 180
XX Db 3935 TGTACAGCTTCCGAGAGAGATACATTCACCTGGAATACTCCACAGATAAAG 3994
XX
XX 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 240
XX Db 3995 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGAATGCGCT 4054
XX
XX 241 GACTCAAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
XX Db 4055 GACTCAAGAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAG 4114
XX
XX 301 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACAGAGAGAGGTCGAATTG 360
XX

```

```

Db 4115 ATAGAAGACTCAGATCTTACATCTGTGAAGTGAGAGACAGAGAGAGGTCGAATTG 4174
Qy 361 CTAGTGTTCGATTTGACCTGACCACTCTGACACCCACTGCTTACAGGGGACGCTGACC 420
Db 4175 CTAGTGTTCGATTTGACCTGACCACTCTGACACCCACTGCTTACAGGGGACGCTGACC 4234
Qy 421 CTGACCTTGGAGAGCCCCCTGTAGTAGTACCCCTCAGTGCAATGTAGAGTCCAGGGGT 480
Db 4235 CTGACCTTGGAGAGCCCCCTGTAGTAGTACCCCTCAGTGCAATGTAGAGTCCAGGGGT 4294
Qy 481 AAAAAATATACAGAGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 540
Db 4295 AAAAAATATACAGAGGGGAGAGACCTCTCCGTGTCTCAGCTGAGCTCCAGATAGTGC 4354
Qy 541 ACCTGACATGACCTGTCTTGTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCTG 600
Db 4355 ACCTGACATGACCTGTCTTGTGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCTG 4414
Qy 601 GTGCTAGCTTTCAGAGGCTCCG 625
Db 4415 GTGCTAGCTTTCAGAGGCTCCG 4439

RESULT 3
AAF82582
ID AAF82582 standard; cDNA, 1377 BP.
AC AAF82582;
XX
XX 18-JUN-2001 (first entry)
XX
XX Human CD4 gene T4.
XX
XX Human; CD4; T4; CD4 fusion protein; oligomerisation;
XX receptor-ligand interaction inhibition; surface plasmon resonance; SPR;
XX T cell receptor binding; MHC binding; carcinoma; autoimmune disease;
XX multiple sclerosis; human immunodeficiency virus; HIV; diabetes;
XX rheumatoid arthritis; immune disorder; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 1..1377
XX FT /*tag= a
XX FT /product= "human CD4"
XX FT sig_peptide 1..75
XX FT /*tag= b
XX FT mat_peptide 76..1374
XX FT /*tag= c
XX
XX WO200122084-A2.
XX
XX 29-MAR-2001.
XX
XX 18-SEP-2000; 2000WO-GB003579.
XX
XX 21-SEP-1999; 99GB-00022352.
XX
XX (AVID-) AVIDEX LTD.
XX
XX Jakobsen BK;
XX
XX WPI; 2001-273470/28.
XX P-PSDB; AAB81502.
XX
XX Sequential screening of candidate compounds library for those which
XX inhibit binding of low affinity receptor-ligand interaction having fast
XX binding kinetics, using interfacial optical assay.
XX
XX Disclosure; Fig 13; 91pp; English.
XX
XX The present sequence encodes human CD4. Human CD4 extracellular domains 1
XX and 2 were used in the construction of CD4 oligomerisation fusion
XX

```





```

RESULT 6
AAQ38760
ID AAQ38760 standard; DNA; 1421 BP.
XX
XX AAQ38760;
AC
XX 25-MAR-2003 (revised)
DT 22-JUL-1993 (first entry)
XX
XX sCD4-L1 lysosomal targeting fusion gene.
DE
XX Soluble CD4; HIV; Human immunodeficiency virus; envelope; glycoprotein;
KM polymerase chain reaction; lysosomal membrane protein; lamp-1; ss.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 1..6
FT /*tag= a
FT /note= "EcoRI restriction site"
FT misc_feature 1275..1280
FT /*tag= b
FT /note= "XbaI restriction site"
FT misc_feature 1416..1421
FT /*tag= c
FT /note= "Sali restriction site"
XX
XX WO9306216-A1.
XX
XX 01-APR-1993.
XX
XX 22-SEP-1992; 92WO-US008090.
XX
XX 26-SEP-1991; 91US-00766963.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX Tang JUN, Lin XL;
XX
XX WPI; 1993-117537/14.
XX
XX New fusion protein used in gene therapy for treating AIDS - comprises
PT protein which binds to retroviral envelope protein which targets fusion
PT protein to lysosome.
XX
XX
XX Claim 12; Page 13-14; 47pp; English.
XX
XX Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI
CC -XbaI fragment of plasmid pT48 containing the sCD4 sequence. The PCR
CC product was ligated to an XbaI-Sali fragment containing the sequence
CC encoding lysosomal membrane protein (LAMP-1) (PCR amplified from a human
CC liver lambda gpi0 library using primers AAQ38754 and AAQ38755). The
CC resulting fusion sequence sCD4-L1 can be used to interfere with the
CC normal function of HIV and to direct the virus to lysosomes. (Updated on
CC 25-MAR-2003 to correct PN field.)
XX
XX
XX Sequence 1421 BP; 355 A; 382 C; 401 G; 283 T; 0 U; 0 Other;
SQ
Query Match 43.4%; Score 624.4; DB 2; Length 1421;
Best Local Similarity 99.8%; Pred. No. 1.6e-129;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 ATGAACCGGGAGAGCCCTTTAGGCACTTGCTTGAGGCTGCAACTGGGCTCTCCCA 60
Db 81 ATGAACCGGGAGAGCTCTTTAGGCACTTGCTTGAGGCTGCAACTGGGCTCTCCCA 140
QY 61 GCAAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGAGTACAGTGAACCTGACC 120
Db 141 GCAAGCACTCAGGAGAAAGAGTGTGCTGGCAAAAAAGGGAGTACAGTGAACCTGACC 200
QY 121 TGTACAGCTTCCCAAGAGAGATACATTTCCACTGGAAAAAATCTCCACAGATTAAG 180
Db 201 TGTACAGCTTCCCAAGAGAGATACATTTCCACTGGAAAAAATCTCCACAGATTAAG 260

```

```

QY 181 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATCGCGCT 240
Db 261 ATTCTGGAAATCAGGGCTCTCTTTAACTAAAGTCCATCCAGCTGATATCGCGCT 320
QY 241 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGTATCATCAAGATCTTTAG 300
Db 321 GACTCAAGAAAGAGCTTTGGGACCAAGAAATTTCCCTGTATCATCAAGATCTTTAG 380
QY 301 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 360
Db 381 ATGAAGACTCAGATCTTACATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 440
QY 361 CTAGTGTTCGATTTGACTGCACTGCACTGACACCCACTGCTTGAAGGAGCCTGACC 420
Db 441 CTAGTGTTCGATTTGACTGCACTGCACTGACACCCACTGCTTGAAGGAGCCTGACC 500
QY 421 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTTCACTGCAATGTAGAGTCCAAAGGCT 480
Db 501 CTGACCTTGGAGAGCCCCCTGGTAGTAGCCCTTCACTGCAATGTAGAGTCCAAAGGCT 560
QY 481 AAAAACAATACAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGGCTCCAGATATGCG 540
Db 561 AAAAACAATACAGGGGGGGAAGACCTCTCCGTCTCAGCTGAGGCTCCAGATATGCG 620
QY 541 ACCTGACATGACATGCTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATGACATCTG 600
Db 621 ACCTGACATGACATGCTCTTGCAGAAACCAAGAAAGTGAAGTTCAAAATGACATCTG 680
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
Db 681 GTGCTAGCTTTCCAGAAAGGCTCCAG 706

RESULT 7
AAQ38759
ID AAQ38759 standard; DNA; 1448 BP.
XX
XX AAQ38759;
AC
XX 25-MAR-2003 (revised)
DT 22-JUL-1993 (first entry)
XX
XX sCD4-HAP lysosomal targeting fusion gene.
DE
XX Soluble CD4; HIV; Human immunodeficiency virus; envelope; ss;
KM glycoprotein; polymerase chain reaction; human acid phosphatase.
XX
XX Synthetic.
OS
XX
XX Key Location/Qualifiers
FH misc_feature 1..6
FT /*tag= a
FT /note= "EcoRI restriction site"
FT misc_feature 1444..1448
FT /*tag= b
FT /note= "Sali restriction site"
XX
XX WO9306216-A1.
XX
XX 01-APR-1993.
XX
XX 22-SEP-1992; 92WO-US008090.
XX
XX 26-SEP-1991; 91US-00766963.
XX
XX (OKLA-) OKLAHOMA MED RES FOUND.
XX
XX Tang JUN, Lin XL;
XX
XX WPI; 1993-117537/14.
XX
XX New fusion protein used in gene therapy for treating AIDS - comprises
PT

```

PT protein which binds to retroviral envelope protein which targets fusion  
 PT protein to lysosome.

PS Claim 12, Page 12-13; 47pp; English.

CC Primers P-1 and P-2 (AAQ38748 and AAQ38749) were used to amplify an EcoRI  
 CC -XbaI fragment of plasmid pT48 containing the scD4 sequence. The PCR  
 CC product was ligated to an XbaI-SalI fragment containing the sequence  
 CC encoding human acid phosphatase (HAP) lysosomal targeting protein (PCR  
 CC amplified from a human liver lambda gp10 library using primers AAQ38752  
 CC and AAQ38753). The resulting fusion sequence scD4-HAP can be used to  
 CC interfere with the normal function of HIV and to direct the virus to  
 CC lysosomes. (Updated on 25-MAR-2003 to correct PN field.)

SQ Sequence 1448 BP; 357 A; 395 C; 404 G; 292 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 2; Length 1448;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 60
DB 81 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 140
OY 61 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACGACC 120
DB 141 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACGACC 200
OY 121 TGTACAGCTTCCAGAGAGAGATACATTCCTGGAATAAAGCTCCACAGATTAAG 180
DB 201 TGTACAGCTTCCAGAGAGAGATACATTCCTGGAATAAAGCTCCACAGATTAAG 260
OY 181 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTCCATCCAGCTGAATATGCGCT 240
DB 261 ATTCTGGGAATCAGGGCTCTTCTTAATAAGTCCATCCAGCTGAATATGCGCT 320
OY 241 GACTCAAGAGAGAGCTTTGGGACCAAGGAACTCCCTGATATCAAGAACTTAG 300
DB 321 GACTCAAGAGAGAGCTTTGGGACCAAGGAACTCCCTGATATCAAGAACTTAG 380
OY 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 360
DB 381 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGTGCATTTG 440
OY 361 CTAGTGTTCGATTAATCACTCTGACACCTGCTTCAAGGGGAGAGCTTGACC 420
DB 441 CTAGTGTTCGATTAATCACTCTGACACCTGCTTCAAGGGGAGAGCTTGACC 500
OY 421 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGTAAGAGTCAAGGGGT 480
DB 501 CTGACCTTGGAGAGCCCCCTGTGTAGTACCCCTCAGTGCATATGTAAGAGTCAAGGGGT 560
OY 481 AAAAATCATACAGAGGGGAGAGAGCCCTCTGCTGCTCAGCTGAGCTCCAGATATGTGC 540
DB 561 AAAAATCATACAGAGGGGAGAGAGCCCTCTGCTGCTCAGCTGAGCTCCAGATATGTGC 620
OY 541 ACCTGAGACATGACATCTTGTGCAAGAACCAAGAGTGAAGTTCAAAATAGACATCGTG 600
DB 621 ACCTGAGACATGACATCTTGTGCAAGAACCAAGAGTGAAGTTCAAAATAGACATCGTG 680
OY 601 GTGCTAGCTTTCCAGAGAGGCTCTCCGG 626
DB 681 GTGCTAGCTTTCCAGAGAGGCTCTCCAG 706

```

RESULT 8  
 AAA50662  
 ID AAA50662 standard; DNA; 1714 BP.

AC AAA50662;

DT 09-JAN-2001 (first entry)

XX

DE DNA encoding CD4-IgM fusion protein CH4Mmu.

XX CD4, IGM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
 KW therapy; diagnosis; ds.

XX Homo sapiens.

OS Location/Qualifiers

FT exon

FT Key

FT exon

FT CDS

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

FT FT

Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

OY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 60
DB 111 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTGTCGCAACTGGCGCTCTCCCA 170
OY 61 GGAGCCACTCAGGGAAGAAAGTGTGCTGGGCAAAAAAGGGGATACAGTGAACGACC 120

```

```
Db 171 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 230
Qy 121 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGAGAAAAAATCCAAACAGATAAAG 180
Db 231 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGAGAAAAAATCCAAACAGATAAAG 290
Qy 181 ATTCTGGGAATCAGGGCTCTCTTCTTAAGTAAAGTCCATCCAGTGAATGATCGGCT 240
Db 291 ATTCTGGGAATCAGGGCTCTCTTCTTAAGTAAAGTCCATCCAGTGAATGATCGGCT 350
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300
Db 351 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 410
Qy 301 ATGAGAACTCAATATTCTATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTTG 360
Db 411 ATGAGAACTCAATATTCTATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTTG 470
Qy 361 CTAGTGTTCGGAATTGACTGCCTGACCACTGACACCCAGCTTTCAGGGGACAGGCTGACC 420
Db 471 CTAGTGTTCGGAATTGACTGCCTGACCACTGACACCCAGCTTTCAGGGGACAGGCTGACC 530
Qy 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGAGTCCAAAGGGGT 480
Db 531 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGAGTCCAAAGGGGT 590
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCTCGTGTCTACAGTGAAGTCCAGAAATGTGGC 540
Db 591 AAAAACAATACAGGGGGGGAAGACCTCTCTCGTGTCTACAGTGAAGTCCAGAAATGTGGC 650
Qy 541 ACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAAGTCAAAATAGACATGTG 600
Db 651 ACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAAGTCAAAATAGACATGTG 710
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCTCGG 626
Db 711 GTGCTAGCTTTCCAGAAAGGCTCTCAG 736

RESULT 9
AAZ4063
ID AAZ44063 standard; DNA; 1714 BP.
XX
AC AAZ44063;
XX
DT 23-MAR-2000 (first entry)
XX
DE Human fusion protein CD4Mg encoding DNA.
XX
KM Fusion protein; human; CD4; IGM; immunoglobulin; gp120;
anti-human immunodeficiency virus; CD4Mg; ds.
XX
OS Homo sapiens.
XX
OS Synthetic.
XX
PN US6004781-A.
XX
PD 21-DEC-1999.
XX
PF 04-FEB-1994; 94US-00191708.
XX
PR 22-JAN-1988; 88US-00147351.
PR 23-JAN-1989; 89US-00299596.
PR 09-JUN-1992; 92US-00896781.
PR 12-APR-1993; 93US-00057952.
XX
PA (GEHO ) GEN HOSPITAL CORP.
XX
PI Seed B;
XX
WPI; 2000-085792/07.
XX
P-PSDB; AAY51080.
```

```
XX
PT Fusion protein useful for the treatment of human immunodeficiency virus.
XX
PS Example 1; Col 41-50; 39pp; English.
XX
CC This invention describes a novel nucleic acid (I) encoding a fusion
CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
CC chain (III). The products of the invention have anti-human
CC immunodeficiency virus (HIV) activity and are capable of binding to
CC gp120. The fusion protein is useful for treating human immunodeficiency
CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence encodes
CC the fusion protein CD4Mg which is constructed from CD4 linked to human
CC IGM upstream of the CH1 region
XX
SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;
Query Match 43.4%; Score 624.4; DB 3; Length 1714;
Best Local Similarity 99.8%; Pred. No. 1,6e-129;
Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGTGCACTGGGCTCCCTCCA 60
Db 111 ATGACCGGGGAGTCCCTTTAGGCACTTGCTGTGTGCACTGGGCTCCCTCCA 170
Qy 61 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 120
Db 171 GCAGCCACTCAGGGAAGAAAGTGTCTGTGGGCAAAAAAGGGGATACAGTGAACCTGACC 230
Qy 121 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGAGAAAAAATCCAAACAGATAAAG 180
Db 231 TGTACAGCTTCCCAAGAAAGAGATACATTCACCTGAGAAAAAATCCAAACAGATAAAG 290
Qy 181 ATTCTGGGAATCAGGGCTCTCTTCTTAAGTAAAGTCCATCCAGTGAATGATCGGCT 240
Db 291 ATTCTGGGAATCAGGGCTCTCTTCTTAAGTAAAGTCCATCCAGTGAATGATCGGCT 350
Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 300
Db 351 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGAAATCTTAAG 410
Qy 301 ATGAGAACTCAATATTCTATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTTG 360
Db 411 ATGAGAACTCAATATTCTATCTGTGAAAGTGAAGACCAAGAAAGAGAGTGCATTTG 470
Qy 361 CTAGTGTTCGGAATTGACTGCCTGACCACTGACACCCAGCTTTCAGGGGACAGGCTGACC 420
Db 471 CTAGTGTTCGGAATTGACTGCCTGACCACTGACACCCAGCTTTCAGGGGACAGGCTGACC 530
Qy 421 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGAGTCCAAAGGGGT 480
Db 531 CTGACCTTGGAGAGCCCCCTGTGATGAGCCCTCACTGCAATGTAGAGAGTCCAAAGGGGT 590
Qy 481 AAAAACAATACAGGGGGGGAAGACCTCTCTCGTGTCTACAGTGAAGTCCAGAAATGTGGC 540
Db 591 AAAAACAATACAGGGGGGGAAGACCTCTCTCGTGTCTACAGTGAAGTCCAGAAATGTGGC 650
Qy 541 ACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAAGTCAAAATAGACATGTG 600
Db 651 ACCTGACATGCACTGTCTTGCAAGAACAGAAAGAGTGAAGTCAAAATAGACATGTG 710
Qy 601 GTGCTAGCTTTCCAGAAAGGCTCTCGG 626
Db 711 GTGCTAGCTTTCCAGAAAGGCTCTCAG 736

RESULT 10
AAZ48203
ID AAZ48203 standard; DNA; 1714 BP.
XX
AC AAZ48203;
XX
DT 14-MAR-2000 (first entry)
```

XX DNA sequence encoding CD4-Ig fusion protein CD4Mmu.  
 DE HIV, extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
 KM secreted protein; HIV infection; medicament; ds.  
 XX Synthetic.  
 OS Homo sapiens.  
 XX CA1340741-C.  
 PN 14-SEP-1999.  
 PD 20-JAN-1989; 89CA-00588749.  
 XX 20-JAN-1989; 89CA-00588749.  
 XX 20-JAN-1989; 89CA-00588749.  
 PR (GENE) GEN HOSPITAL CORP.  
 XX  
 XX Seed B;  
 PI WPI, 2000-063015/06.  
 DR P-PSDB; AAY59170.  
 XX  
 XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
 PT the treatment of HIV or simian immunodeficiency virus infections.  
 XX  
 PS Example 1, Page 47-53, 89pp; English.  
 XX  
 CC The invention provides a fusion gene encoding a fusion protein that  
 CC comprises an extracellular CD4 DNA sequence or its fragment which binds  
 CC to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
 CC sequence of an Ig heavy or light chain, where the DNA sequence encoding  
 CC the variable region has been replaced with the DNA sequence which encodes  
 CC extracellular CD4 or its gp120 binding fragment. The fusion protein is  
 CC capable of being secreted. The fusion proteins are useful for treating  
 CC HIV or SIV infections in animals, preferably humans. They are also useful  
 CC for producing medicaments which can be used for treating HIV or SIV  
 CC infections in humans. The present sequence represents the DNA which  
 CC encodes the fusion protein CD4Mmu where the CD4 is linked to human IgG1  
 CC at the Mst2 site upstream of the CH1 region  
 CC  
 SQ Sequence 1714 BP; 411 A; 503 C; 467 G; 333 T; 0 U; 0 Other;  
 Query Match 43.4%; Score 624.4; DB 3; Length 1714;  
 Best Local Similarity 99.8%; Pred. No. 1.6e-129;  
 Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTCGCACTGGCGCTCCCA 60  
 DB 111 ATGAACCGGGAGTCCCTTTAGGCACTTGTGTCGCACTGGCGCTCCCA 170  
 QY 61 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCAAAAAAGGATACAGTGAAC 120  
 DB 171 GCAGCCACTCAGGGAAGAAAGTGTGTGTCGCAAAAAAGGATACAGTGAAC 230  
 QY 121 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATAAAGTCCCAAGATAAG 180  
 DB 221 TGTACAGCTTCCCAAGAAAGCATACATTCCTGGAATAAAGTCCCAAGATAAG 290  
 QY 181 ATTCGGGGAATCAGGGCTCCTTTAAGTAAAGTCCATCCAACTGAATATGGCGCT 240  
 DB 291 ATTCGGGGAATCAGGGCTCCTTTAAGTAAAGTCCATCCAACTGAATATGGCGCT 350  
 QY 241 GACTCAAGAGAGGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300  
 DB 351 GACTCAAGAGAGGCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 410  
 QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360  
 DB 411 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 470  
 QY 361 CTAGTGTTCGATGACTGCCAACTGTGACACCACTGCTTCAGGGGAGAGGCTGACC 420

DB 471 CTAGTGTTCGATGACTGCCAACTGTGACACCACTGCTTCAGGGGAGAGGCTGACC 530  
 QY 421 CTGACCTTGAGAGAGCCCCCTGTAGTAGAGCCCTCAGTGCATGTAGAGATCCAGGGGT 480  
 DB 531 CTGACCTTGAGAGAGCCCCCTGTAGTAGAGCCCTCAGTGCATGTAGAGATCCAGGGGT 590  
 QY 481 AAAACATACAGGGGGGGAAGACCTCTCCGTCGTACGTGAGCTCCAGATATGTGAC 540  
 DB 591 AAAACATACAGGGGGGGAAGACCTCTCCGTCGTACGTGAGCTCCAGATATGTGAC 650  
 QY 541 ACCTGACATGACATGCTGTCTTGCAGAACCAAGAGTGTGAGTCAAAATAGACATTCGTG 600  
 DB 651 ACCTGACATGACATGCTGTCTTGCAGAACCAAGAGTGTGAGTCAAAATAGACATTCGTG 710  
 QY 601 GTGCTAGCTTTCAGAGAGGCTCCGG 626  
 DB 711 GTGCTAGCTTTCAGAGAGGCTCCAG 736  
 RESULT 11  
 AAA35205  
 ID AAA35205 standard; DNA; 1742 BP.  
 XX  
 XX AAA35205;  
 AC  
 XX 28-JUL-2000 (first entry)  
 DT  
 XX  
 DE Human adenosine receptor related polynucleotide 2nd SEQ ID NO:79.  
 XX  
 KW Human; adenosine receptor; low adenosine antisense oligonucleotide;  
 KW phosphorothioate; impaired respiration; inflammation; allergy;  
 KW allergic disease; bronchoconstriction; inhibitor; antiinflammatory;  
 KW antiallergic; anticholinergic; cytosolic; analgesic; impaired airway;  
 KW lung disease; ischaemic condition; pulmonary vasoconstriction; asthma;  
 KW respiratory distress syndrome; pain; cystic fibrosis; emphysema;  
 KW pulmonary hypertension; chronic obstructive pulmonary disease; COPD;  
 KW cancer; leukaemia; lymphoma; carcinoma; metastasis; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200009525-A2.  
 XX  
 PD 24-FEB-2000.  
 XX  
 XX 03-AUG-1999; 99WO-US017712.  
 PF  
 XX 03-AUG-1999; 98US-0095212P.  
 PR  
 XX (UYEC-) UNIV EAST CAROLINA.  
 PA  
 XX Nyce JW;  
 PI  
 XX WPI, 2000-205971/18.  
 DR  
 XX  
 PT New antisense oligonucleotides useful for treating e.g. pulmonary  
 PT vasoconstriction, inflammation, allergies, asthma, hypertension,  
 PT bronchitis, emphysema, respiratory distress syndrome, ischemia or  
 PT cancers.  
 PT  
 XX  
 PS Disclosure: Page 1245, 1343pp; English.  
 XX  
 CC The present invention describes a new composition comprising an antisense  
 CC oligonucleotide (ON) with low adenosine (up to 15%), which targets  
 CC nucleic acids involved in bronchoconstriction, allergies, and/or  
 CC inflammation. The ON can have antiinflammatory, antiallergic,  
 CC anticholinergic, cytosolic and analgesic activities. The compositions are  
 CC useful for the treatment of diseases associated with inflammation,  
 CC impaired airways, including lung disease and diseases whose secondary  
 CC effects afflict the lungs of a subject. They can be used for treating  
 CC e.g. ischaemic conditions, pulmonary vasoconstriction, allergies, asthma,  
 CC impaired respiration, respiratory distress syndrome, pain, cystic  
 CC fibrosis, pulmonary hypertension, emphysema, chronic obstructive

CC pulmonary disease (COPD), and cancers such as leukaemias, lymphomas,  
 CC carcinomas, and cancers which may metastasize to the lungs, including  
 CC breast and prostate cancer. The reduction of the adenosine content of the  
 CC ONS reduces side effects. The A-containing ONS break down with the  
 CC release of deoxyadenosine which activates adenosine receptors causing the  
 CC bronchoconstriction and inflammation. AAA3313 to AAA3512 represent the  
 CC nucleotide sequences given in the sequence listing from the present  
 CC invention, which correspond to SEQ ID NO:1 to 2815, and then the last 185  
 CC sequences are also called SEQ ID NO:1 to 185, but the sequences differ  
 CC from the previously named sequences. SEQ ID NO:11 to 1680 (AAA3233 to  
 CC AAA3392) are specifically claimed ONS from the present invention. N.B.  
 CC Sequences given in the disclosure of the present invention do not match  
 CC up with their corresponding SEQ ID NO: sequences given in the sequence  
 CC listing

XX  
 XX Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 60  
 Db 76 ATGAAACCGGGAGTCCCTTTTAAAGCACTTGTGTGCTGCAACTGGCGCTCTCCCA 135  
 Qy 61 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 120  
 Db 136 GCAGCCACTCAGGGAAGAAAGTGTGCTGGCAAAAAAGGGATACAGTGAATGACC 195  
 Qy 121 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAAGAACTCCACACAGATAAG 180  
 Db 196 TGTACAGCTTCCAGAAAGAGCATATTCACCTGGAAGAACTCCACACAGATAAG 255  
 Qy 181 ATTCTGGGAATCGGGCTCTTCTTAACTGAAGTCAATCCAGCTGAATGATCGGCT 240  
 Db 256 ATTCTGGGAATCGGGCTCTTCTTAACTGAAGTCAATCCAGCTGAATGATCGGCT 315  
 Qy 241 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 300  
 Db 316 GACTCAAGAAAGACCTTTGGGACCAAGAAATCTCCCTGATCATCAAGATCTTAAG 375  
 Qy 301 ATAGAAAGCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 360  
 Db 376 ATAGAAAGCTCAGTACTTACATCTGTGAAGTGAAGACCAAGAAAGAGGTGCAATTG 435  
 Qy 361 CTAGTGTTCGATTTGATGCTGCAACTGACCCCACTGCTTCAAGGGCAAGCTTACC 420  
 Db 436 CTAGTGTTCGATTTGATGCTGCAACTGACCCCACTGCTTCAAGGGCAAGCTTACC 495  
 Qy 421 CTGACCTTGGAGAGCCCTGAGTAGACCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 480  
 Db 496 CTGACCTTGGAGAGCCCTGAGTAGAGCCCTCAGTGCAATGTAGAGTCCAAAGGGGT 555  
 Qy 481 AAAAATACAGAGGGGGGAGAAACCTCTCTGCTGTCTCAGCTGAGCTCCAGATAGTGC 540  
 Db 556 AAAAATACAGAGGGGGGAGAAACCTCTCTGCTGTCTCAGCTGAGCTCCAGATAGTGC 615  
 Qy 541 ACCGTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 600  
 Db 616 ACCGTGACATGCACTGTCTTGCAGAACCAAGAAAGGTGAGTTCAAAATAGACATGCTG 675  
 Qy 601 GTGCTAGCTTCCAGAGAGCTCCGG 626  
 Db 676 GTGCTAGCTTCCAGAGAGCTCCAG 701

RESULT 12

AAF21327

ID AAF21327 standard; DNA; 1742 BP.

XX AAF21327;

XX 14-MAR-2001 (first entry)

XX  
 XX Human low adenosine antisense oligonucleotide related sequence #2894.

DE Low adenosine antisense oligonucleotide; phosphorothioate; allergy;  
 XX human; airway disorder; bronchoconstriction; lung inflammation;  
 KW surfactant depletion; respiratory; bronchodilator; antiinflammatory;  
 KW immunosuppressive; antiaesthetic; analgesic; hypotensive; cytostatic;  
 KW respiratory obstruction; pulmonary obstruction; impeded respiration;  
 KW surfactant hypoproduction; pulmonary vasoconstriction; asthma; RDS;  
 KW respiratory distress syndrome; pain; cystic fibrosis; allergic rhinitis;  
 KW pulmonary hypertension; emphysema; pulmonary transplantation rejection;  
 KW chronic obstructive pulmonary disease; pulmonary infection; bronchitis;  
 KW cancer; ss.

XX Homo sapiens.

XX W0200062736-A2.

XX 26-OCT-2000.

XX 24-MAR-2000; 2000MO-US008020.

XX 06-APR-1999; 99US-0127958P.

XX (UYEC-) UNIV EAST CAROLINA.

XX (NYCE/) NYCE J W.

XX Nyce JW;

XX WPI; 2000-679539/66.

XX Low adenosine (A) content antisense oligonucleotides which do not trigger  
 PT adenosine receptors during metabolism, useful e.g. for treating cancers  
 PT and respiratory obstructions.

PS Disclosure, Page 1329; 1592pp; English.

XX The present invention describes low adenosine (A) content antisense  
 CC oligonucleotides and compositions (I) comprising them. In the antisense  
 CC oligonucleotides the A is replaced by a 'Universal' or alternative base.  
 CC (I) can have respiratory, bronchodilator, antiinflammatory, analgesic,  
 CC immunosuppressive, antiaesthetic, hypotensive and cytostatic activities.  
 CC The antisense oligonucleotides and (I) can be used to down-regulate the  
 CC expression and/or activity of target polypeptides associated with  
 CC lung/respiratory disorders and malignancies, such as stimulating and  
 CC activating peptide factors and transmitters, transcription factors,  
 CC immunoglobulins and antibodies, antibody receptors, cytokines and  
 CC chemokines, endogenously produced specific and non-specific enzymes,  
 CC binding proteins, adhesion molecules and their receptors, cytokines and  
 CC chemokine receptors, adenosine receptors, bradykinin receptors, central  
 CC nervous system (CNS) and peripheral nervous and non-nervous system  
 CC receptors, CNS and peripheral nervous and non-nervous system peptide  
 CC transmitters, defensins, growth factors, vasoactive peptides and  
 CC receptors, binding proteins and malignancy associated proteins. The  
 CC antisense oligonucleotides may be used in this way to treat disorders  
 CC including respiratory obstruction (especially pulmonary obstruction  
 CC and/or bronchoconstriction) and/or lung inflammation, allergy(ies) and/or  
 CC surfactant hypoproduction which are associated with a disease or  
 CC condition selected from pulmonary vasoconstriction, inflammation,  
 CC allergies, asthma, impeded respiration, respiratory distress syndrome  
 CC (RDS), pain, cystic fibrosis (CF), allergic rhinitis (AR), pulmonary  
 CC hypertension, emphysema, chronic obstructive pulmonary disease (COPD),  
 CC pulmonary transplantation rejection, pulmonary infections, bronchitis,  
 CC and/or cancer. AAF18434 to AAF21543 represent human polynucleotide  
 CC fragments and antisense oligonucleotides used in the exemplification of  
 CC the present invention

XX Sequence 1742 BP; 405 A; 492 C; 466 G; 379 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 3; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCCCTCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCCCTCCA 135
QY 61 GCAGCCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAAGTAC 120
DB 136 GCAGCCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAAGTAC 195
QY 121 TGTACAGCTCCCAAGAAAGAGCATACATTCCTGAGAAAAAATCCACCAAGATTAAG 180
DB 196 TGTACAGCTCCCAAGAAAGAGCATACATTCCTGAGAAAAAATCCACCAAGATTAAG 255
QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGGTGAATGATGCGCT 240
DB 256 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGGTGAATGATGCGCT 315
QY 241 GACTCAAGAAAGCCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGCCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGTCAATTG 360
DB 376 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGTCAATTG 435
QY 361 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 420
DB 436 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 495
QY 421 CTGACCTTGGAGAGCCCCCTGTTAGTAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTTAGTAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 555
QY 461 AAAAATCATACAGGGGGGAGAGACCTCTCTGCTGTCTGAGTGAAGTCCAGATGTGGC 540
DB 556 AAAAATCATACAGGGGGGAGAGACCTCTCTGCTGTCTGAGTGAAGTCCAGATGTGGC 615
QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
DB 616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

```

## RESULT 13

AA165462  
AA165462 standard; DNA; 1742 BP.

AC AA165462;

DT 10-DEC-2001 (first entry)

XX Nucleotide sequence of a human polynucleotide.

KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;

KM C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; ss.

OS Homo sapiens.

PN MO200164752-A2.

PD 07-SEP-2001.

PE 28-FEB-2001; 2001WO-US006322.

PR 02-MAR-2000; 2000US-00517605.

XX (UNY ) UNIV NEW YORK STATE.

XX (UNY-) UNIV NIJMEGEN.

XX Litman DR, Kwon D, Van Kooyk Y, Geijtenbeek T;

DR WPI; 2001-602565/68.

XX An antibody for the treatment or prevention of HIV-infection comprises a  
PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
PT DC-SIGN due to concomitant conformational change.

PS Disclosure; Page 121-122; 131pp; English.

XX The specification describes an antibody which is specific for an  
CC antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
CC is exposed upon gp120 binding of DC-SIGN due to concomitant  
CC conformational change. DC-SIGN is a receptor that is specifically  
CC expressed on dendritic cells and facilitates infection of T lymphocytes  
CC with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
CC -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
CC with high affinity. The antibody of the invention inhibits the trans  
CC enhancement of HIV entry into a T cell or macrophage facilitated by  
CC dendritic cells. The antibody is useful to treat or prevent HIV  
CC infection. The present sequence represents a human polynucleotide, which  
CC is used in the course of the invention.

SQ Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 4; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129; Mismatches 1; Indels 0; Gaps 0;

```

QY 1 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCCCTCCA 60
DB 76 ATGAACCGGGAGTCCCTTTAGGCACTTGCTTGTGCTGCAACTGGCGCTCCCTCCA 135
QY 61 GCAGCCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAAGTAC 120
DB 136 GCAGCCACTCAGGGAGAAAGAGTGTGCTGGGCAAAAAAGGGGATACAGTGAAGTAC 195
QY 121 TGTACAGCTCCCAAGAAAGAGCATACATTCCTGAGAAAAAATCCACCAAGATTAAG 180
DB 196 TGTACAGCTCCCAAGAAAGAGCATACATTCCTGAGAAAAAATCCACCAAGATTAAG 255
QY 181 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGGTGAATGATGCGCT 240
DB 256 ATTCTGGAAATCAGGGCTCTCTTAACTAAAGGTCATCCAGGTGAATGATGCGCT 315
QY 241 GACTCAAGAAAGCCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 300
DB 316 GACTCAAGAAAGCCCTTTGGAGCAAGAACTTCCCTGATCATCAAGATCTTAAG 375
QY 301 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGTCAATTG 360
DB 376 ATGAAGACTCAGATACCTTACATCTGTGAAGTGAAGAGACCAAGAGAGAGTCAATTG 435
QY 361 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 420
DB 436 CTAGTGTGAGATTGACTGCACTGCACTGCACTGCACTGCACTGCACTGCACTGCACT 495
QY 421 CTGACCTTGGAGAGCCCCCTGTTAGTAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 480
DB 496 CTGACCTTGGAGAGCCCCCTGTTAGTAGCCCCCTGATGCAATGTAGAGTCCAGGGGT 555
QY 461 AAAAATCATACAGGGGGGAGAGACCTCTCTGCTGTCTGAGTGAAGTCCAGATGTGGC 540
DB 556 AAAAATCATACAGGGGGGAGAGACCTCTCTGCTGTCTGAGTGAAGTCCAGATGTGGC 615
QY 541 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 600
DB 616 ACCTGACATGCACTGTCTTGGAGAACCAAGAAAGGTGAGTTCAAAATAGACATCGTG 675
QY 601 GTGCTAGCTTTCCAGAAAGGCTCCGG 626
DB 676 GTGCTAGCTTTCCAGAAAGGCTCCAG 701

```

## RESULT 14



PT diagnosing and treating an immunopathology, e.g. Crohn's disease, asthma  
PT or osteoarthritis.

XX Claim 1; SEQ ID NO 1013; 50pp; English.

CC The invention relates to a composition comprising a plurality of cDNAs  
CC for detecting the altered expression of genes in an immunological  
CC response. The invention also relates to a method of diagnosing or  
CC monitoring the treatment of an immunopathological condition in a sample,  
CC comprising obtaining nucleic acids from a sample, contacting the nucleic  
CC acids of the sample with an array comprising the plurality of cDNAs under  
CC conditions to form one or more hybridisation complexes, detecting the  
CC hybridisation complexes and comparing the levels of the detected  
CC hybridisation complexes with the level of hybridisation complexes  
CC detected in a non-diseased sample, where an altered level of the detected  
CC hybridisation complexes correlates with the presence of an  
CC immunopathological condition. Also disclosed are an expression profile  
CC comprising a microarray and a plurality of detectable complexes and a  
CC method for identifying a plurality of polynucleotide probes. The cDNAs  
CC are useful as hybridisable array elements in a microarray for monitoring  
CC the expression of target polynucleotides. The microarray can be used in  
CC the diagnosis of an immunopathology, such as Crohn's disease, asthma,  
CC ulcerative colitis, hyperesinophilia, irritable bowel syndrome,  
CC osteoarthritis, rheumatoid arthritis or acute monocytic leukaemia, and in  
CC identifying agents for the treatment of the diseases. The microarray may  
CC also be used in drug discovery and development, toxicological and  
CC carcinogenicity studies, forensics or pharmacogenomics. The composition  
CC may also be used in purification of a subpopulation of mRNAs, cDNAs or  
CC genomic fragments. This sequence data for this patent did not form part of  
CC the printed specification but was obtained in electronic format directly  
CC from USPTO at [seqdata.uspto.gov/sequence.html](http://seqdata.uspto.gov/sequence.html).

XX Sequence 1742 BP; 405 A; 491 C; 466 G; 380 T; 0 U; 0 Other;

Query Match 43.4%; Score 624.4; DB 11; Length 1742;

Best Local Similarity 99.8%; Pred. No. 1.6e-129;

Matches 625; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```
OY 1 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGCACTGGCGCTCTCCCA 60
Db 76 ATGAACCGGGAGTCCCTTTTAGGCACTTGTGTCGTCGCACTGGCGCTCTCCCA 135
OY 61 GCAGCCTCTCAGGGAAGAAAGTGTGTGTCGTCGCACTGGCGCTCTCCCA 120
Db 136 GCAGCCTCTCAGGGAAGAAAGTGTGTGTCGTCGCACTGGCGCTCTCCCA 195
OY 121 TGTACAGCTTCCAGAGAAGAGATACAAATTCACCTGGAAAACTCCAAACAGATTAAG 180
Db 196 TGTACAGCTTCCAGAGAAGAGATACAAATTCACCTGGAAAACTCCAAACAGATTAAG 255
OY 181 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGGTCCATCCAGCTGAATGCGGCT 240
Db 256 ATTCTGGGAATCAGGGCTCTCTTTAACTAAAGGTCCATCCAGCTGAATGCGGCT 315
OY 241 GACTCAAGAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAAG 300
Db 316 GACTCAAGAAGAGCTTTGGGACCAAGAACTTCCCTGATCATCAAGATCTTAAAG 375
OY 301 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 360
Db 376 ATAGAAGACTCAGATTAATCATCTGTGAAGTGAAGACCAAGAGAGAGGTGCAATTG 435
OY 361 CTAGTGTTCGATTAATCACTCTGACCACTGCTTCAGGGGAGAGCTGACC 420
Db 436 CTAGTGTTCGATTAATCACTCTGACCACTGCTTCAGGGGAGAGCTGACC 495
OY 421 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGTCAAGGGGT 480
Db 496 CTGACCTTGAAGAGCCCCCTGTGTAGTACCCCTCAATGCAATGTAAGAGTCAAGGGGT 555
OY 481 AAAAATATACAGGGGGGAGAGCCCTTCCTGCTCAGCTGAGCTCCAGGATAGTGGC 540
```

```
Db 556 AAAAATATACAGGGGGGAGAGCCCTTCCTGCTCAGCTGAGCTCCAGATAGTGGC 615
OY 541 ACTTGACATGCACTGTCTTTGAGAACCAAGAAAGTGTGAGTTCAAAATATGACATCGTG 600
Db 616 ACTTGACATGCACTGTCTTTGAGAACCAAGAAAGTGTGAGTTCAAAATATGACATCGTG 675
OY 601 GTGCTAGCTTCCAGAGAGGCGCTCCGG 626
Db 676 GTGCTAGCTTCCAGAGAGGCGCTCCAG 701
```

Search completed: November 22, 2004, 11:05:33  
Job time : 760 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:56 ; Search time 65.8175 Seconds  
(without alignments)  
511.863 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668

Sequence: 1 MNRGVPRFHLVLVQLALLP.....GTRLEIKLVPRSGHHHHH 508

Scoring table:

BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:\*

1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTCUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1064	39.9	295	6	5223394-9
2	1064	39.9	318	6	5223394-11
3	1064	39.9	458	3	US-09-039-555B-15
4	1064	39.9	458	3	US-09-517-605-3
5	1062	39.8	458	4	US-10-092-138A-25
6	1059	39.7	398	2	US-08-284-391B-29
7	1059	39.7	398	4	US-09-218-950-29
8	1059	39.7	398	4	US-08-394-388A-29
9	1059	39.7	402	1	US-08-236-311-1
10	1059	39.7	402	3	US-08-457-918-1
11	1059	39.7	402	4	US-10-157-408-1
12	1059	39.7	457	4	US-08-328-500-9
13	1059	39.7	458	3	US-08-466-368-4
14	1059	39.7	462	2	US-08-417-495-5
15	1059	39.7	462	2	US-08-284-391B-5
16	1059	39.7	462	3	US-09-210-950-5
17	1059	39.7	462	4	US-08-394-388A-5
18	1059	39.7	462	5	PCT-US92-01785-5
19	1059	39.7	462	5	PCT-US95-00454-5
20	1059	39.7	532	2	US-08-417-495-6
21	1059	39.7	532	2	US-08-284-391B-6
22	1059	39.7	532	4	US-09-218-950-6
23	1059	39.7	532	4	US-08-394-388A-6
24	1059	39.7	532	5	PCT-US92-01785-6
25	1059	39.7	532	5	PCT-US95-00454-6
26	1059	39.7	575	2	US-08-417-495-4
27	1059	39.7	575	2	US-08-284-391B-4

28	1059	39.7	575	3	US-09-218-950-4	Sequence 4, Appl1
29	1059	39.7	575	4	US-08-394-388A-4	Sequence 4, Appl1
30	1059	39.7	575	5	PCT-US92-01785-4	Sequence 4, Appl1
31	1059	39.7	575	5	PCT-US95-00454-4	Sequence 4, Appl1
32	1059	39.7	630	4	US-08-472-888A-6	Sequence 6, Appl1
33	1056	39.6	458	6	5223394-7	Patent No. 5223394
34	1053	39.5	394	4	US-08-466-368-2	Sequence 2, Appl1
35	1053	39.5	394	4	US-08-328-500-2	Sequence 2, Appl1
36	1051	39.4	530	3	US-08-477-460B-4	Sequence 4, Appl1
37	1051	39.4	530	3	US-08-379-516-4	Sequence 4, Appl1
38	1051	39.4	530	3	US-09-329-916-4	Sequence 4, Appl1
39	1051	39.4	530	3	US-08-485-372A-4	Sequence 4, Appl1
40	1051	39.4	530	3	US-09-409-006A-4	Sequence 4, Appl1
41	1051	39.4	530	4	US-08-484-681-4	Sequence 4, Appl1
42	1051	39.4	530	4	US-09-766-995-4	Sequence 4, Appl1
43	1051	39.4	530	5	PCT-US93-07422-4	Sequence 4, Appl1
44	1050	39.4	432	3	US-08-477-460B-2	Sequence 2, Appl1
45	1050	39.4	432	3	US-08-379-516-2	Sequence 2, Appl1

## ALIGNMENTS

RESULT 1  
5223394-9  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO: 9:  
; LENGTH: 295

Query Match 39.9%; Score 1064; DB 6; Length 295;  
Best Local Similarity 100.0%; Pred. No. 8.3e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MNRGVPRFHLVLVQLALLPAAATGKGVVIGKGGDTVELTCTASQKKSIOFMKNSNOIK	60
DB	1	MNRGVPRFHLVLVQLALLPAAATGKGVVIGKGGDTVELTCTASQKKSIOFMKNSNOIK	60
QY	61	ILNGSFLTKGPSKLNDRADSRSLMDQGNFPLIKLIKIEDSDTYICEVEDQKEEYOL	120
DB	61	ILNGSFLTKGPSKLNDRADSRSLMDQGNFPLIKLIKIEDSDTYICEVEDQKEEYOL	120
QY	121	LVFGLTANSDFHLQGGSLITLTLESPGSSPSVQGRSPRGKNIQGGKTLVSQLELDQSG	180
DB	121	LVFGLTANSDFHLQGGSLITLTLESPGSSPSVQGRSPRGKNIQGGKTLVSQLELDQSG	180
QY	181	TWTCVLNOKKVEFKIDIVLAFQKAS	208
DB	181	TWTCVLNOKKVEFKIDIVLAFQKAS	208

RESULT 2  
5223394-11  
; Patent No. 5223394  
; APPLICANT: WALLNER, BARBARA  
; TITLE OF INVENTION: RECOMBINANT DNA MOLECULE COMPRISING  
; LYMPHOCYTE FUNCTION-ASSOCIATED ANTIGEN 3 PHOSPHATIDYLINOSITOL  
; LINKAGE SIGNAL SEQUENCE  
; NUMBER OF SEQUENCES: 12  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/335,688  
; FILING DATE: 10-APR-1989  
; SEQ ID NO: 11:  
; LENGTH: 318

Query Match 39.9%; Score 1064; DB 6; Length 318;  
Best Local Similarity 100.0%; Pred. No. 9e-54;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208

RESULT 3  
US-09-039-555B-15  
; Sequence 15, Application US/09039555B  
; Patent No. 6033856  
; GENERAL INFORMATION:  
; APPLICANT: Koerner, Kathrin  
; APPLICANT: Mueller, Rolf  
; APPLICANT: Sadiacek, Hans-Harald  
; TITLE OF INVENTION: PROMOTER OF THE CDC25B GENE, ITS  
; TITLE OF INVENTION: PREPARATION AND USE  
; NUMBER OF SEQUENCES: 19  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 3000 K Street, N.W., Suite 500  
; CITY: Washington  
; STATE: D. C.  
; COUNTRY: USA  
; ZIP: 20007-5109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/039,555B  
; FILING DATE: 16-MAR-1998  
; CLASSIFICATION: 514  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: DE 19710643.9  
; FILING DATE: 14-MAR-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Bent, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 016779/0131  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 672-5300  
; TELEFAX: (202) 672-5399  
; TELEX: 904136  
; INFORMATION FOR SEQ ID NO: 15:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 458 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-09-039-555B-15

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208

RESULT 4  
US-09-517-605-3  
; Sequence 3, Application US/09517605  
; Patent No. 6391567  
; GENERAL INFORMATION:  
; APPLICANT: Littman, Dan R.  
; APPLICANT: Kwon, Douglas S.  
; APPLICANT: van Kooyk, Yvette  
; APPLICANT: Geijtenbeek, Theo  
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY INTO  
; FILE REFERENCE: 1049-1-017  
; CURRENT APPLICATION NUMBER: US/09/517,605  
; CURRENT FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 458  
; TYPE: PR  
; ORGANISM: Homo sapiens  
US-09-517-605-3

Query Match 39.9%; Score 1064; DB 3; Length 458;  
Best Local Similarity 100.0%; Pred. No. 1.3e-53;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATGKKVVLGKKGDVLELTCTASQKKSIOFHMKNQIK 60  
QY 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
DB 121 LVFGLTANSDTHLLQGOSLTLTLESPPGSSPSVQCSPRGKNIQGGKTLVSQLELDG 180  
QY 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208  
DB 181 TWCTVLQONQKVEFKIDIVLAFQKAS 208

RESULT 5  
US-10-092-138A-25  
; Sequence 25, Application US/10092138A  
; Patent No. 6743630  
; GENERAL INFORMATION:  
; APPLICANT: Sato, Taka-Aki  
; TITLE OF INVENTION: METHOD OF PREPARING A PROTEIN ARRAY BASED ON  
; TITLE OF INVENTION: BIOCHEMICAL PROTEIN-PROTEIN INTERACTION  
; FILE REFERENCE: 65823/jpw/pt  
; CURRENT APPLICATION NUMBER: US/10/092,138A  
; CURRENT FILING DATE: 2002-03-06  
; NUMBER OF SEQ ID NOS: 34  
; SOFTWARE: Patentin Ver. 2.1

SEQ ID NO 25  
LENGTH: 458  
TYPE: PRT  
ORGANISM: human  
US-10-092-138A-25

Query Match 39.8%; Score 1062; DB 4; Length 458;  
Best Local Similarity 99.5%; Pred. No. 1.7e-53;  
Matches 207; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLILVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 NMRGVPFRHLILVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIIKNIKIEDSPYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIIKNIKIEDSPYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDFTHLQGGSLITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFTHLQGGSLITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

## RESULT 6

US-08-284-391B-29  
Sequence 29, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elding LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elding, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-29

Query Match 39.7%; Score 1059; DB 2; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 NMRGVPFRHLILVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
DB 1 NMRGVPFRHLILVLQALLPAATQKRVVLGKKGDVVELTCTASQKKSIOFHMKNSNOIK 60  
QY 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIIKNIKIEDSPYICEVEDQKEEVOL 120  
DB 61 ILGNQGSFLTKGPSKLNDRADSRRLMDQGNPPLIIKNIKIEDSPYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDFTHLQGGSLITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFTHLQGGSLITLTLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLQNKQKVEFKIDIVLAFOKAS 208

## RESULT 7

US-09-218-950-29  
Sequence 29, Application US/09218950  
Patent No. 6284240  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elding LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/218,950  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elding, Karen L  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 29:

SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-218-950-29

Query Match 39.7%; Score 1059; DB 3; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVVLEKKGDVLELTCTASQKSIQFHMKNQIK 60  
Db 1 MNRGVPFRHLLVQLALPAAATQGNKVVLEKKGDVLELTCTASQKSIQFHMKNQIK 60  
QY 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQONQKVEFKIDIVLAFQKAS 208  
Db 181 TWICTVLQONQKVEFKIDIVLAFQKAS 208

RESULT 8  
US-08-394-388A-29  
Sequence 29, Application US/08394388A  
Patent No. 6753162

## GENERAL INFORMATION:

APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
TITLE OF INVENTION: CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Elbing LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/394,388A  
FILING DATE: 24-FEB-1995  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/284,391  
FILING DATE: 02-AUG-1994  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elbing, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045

TELEX:  
INFORMATION FOR SEQ ID NO: 29:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 398 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-394-388A-29

Query Match 39.7%; Score 1059; DB 4; Length 398;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALPAAATQGNKVVLEKKGDVLELTCTASQKSIQFHMKNQIK 60  
Db 1 MNRGVPFRHLLVQLALPAAATQGNKVVLEKKGDVLELTCTASQKSIQFHMKNQIK 60  
QY 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
Db 61 ILGNQGSFLLTKGPKSLNDRADSRSLMDQGNFLLIKNLKIEDSDTYICEVEDQKEEVOL 120  
QY 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDTHLLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWICTVLQONQKVEFKIDIVLAFQKAS 208  
Db 181 TWICTVLQONQKVEFKIDIVLAFQKAS 208

RESULT 9  
US-08-236-111-1  
Sequence 1, Application US/08236311  
Patent No. 556335

## GENERAL INFORMATION:

APPLICANT: Capon, Daniel J.  
APPLICANT: Gregory, Timothy J.  
TITLE OF INVENTION: Adhesion Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94060

## COMPUTER READABLE FORM:

MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,311  
FILING DATE: 02-MAY-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Haack, Janet E.  
REGISTRATION NUMBER: 28,616  
REFERENCE/DOCKET NUMBER: 444PIC2  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 415/225-1896  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-236-311-1

Query Match 39.7%; Score 1059; DB 1; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPRHLLVLTALPAATQGGKGVETCTASQKKSIOFHMKNNOIK 60  
DB 1 MRGVPRHLLVLTALPAATQGGKGVETCTASQKKSIOFHMKNNOIK 60  
QY 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGILTANSDDLTLGGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGILTANSDDLTLGGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208

RESULT 10  
US-08-457-918-1  
Sequence 1, Application US/08457918

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Timothy J.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/457, 918  
FILING DATE: 1-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.

REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168

INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-457-918-1

Query Match 39.7%; Score 1059; DB 3; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPRHLLVLTALPAATQGGKGVETCTASQKKSIOFHMKNNOIK 60  
DB 1 MRGVPRHLLVLTALPAATQGGKGVETCTASQKKSIOFHMKNNOIK 60  
QY 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILGNQSFLLTKGPSKLNDRADSRSLMDQGNPPLIIKLIKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGILTANSDDLTLGGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGILTANSDDLTLGGOSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208  
DB 181 TWTCTVLONQKKVEFKIDIVLAFOKAS 208

RESULT 11

US-10-157-408-1  
Sequence 1, Application US/10157408

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
TITLE OF INVENTION: Timothy J.  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESSES:  
ADDRESSER: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/157, 408  
FILING DATE: 28-MAY-2002  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457, 918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:

NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-10-157-408-1  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
Query Match 39.7%; Score 1059; DB 4; Length 402;  
Best Local Similarity 99.5%; Pred. No. 2.2e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVQLALPAPATQGGKGVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLLVQLALPAPATQGGKGVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILGNQGSFLTQGPSKLNDRADSRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILGNQGSFLTQGPSKLNDRADSRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRKNIOGKKTLSVSOLELDDSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRKNIOGKKTLSVSOLELDDSG 180  
QY 181 TWCTVQLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWCTVQLQNKQKVEFKIDIVLAFQKAS 208  
RESULT 12  
US-08-328-500-9  
Sequence 9, Application US/08328500  
Patent No. 6673896  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Axel, Richard  
APPLICANT: Sweet, Richard W.  
APPLICANT: Arthos, James  
TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
NUMBER OF SEQUENCES: 22  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham  
STREET: 1185 Avenue of the Americas  
CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/328,500  
FILING DATE:  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 0575/24577-CY  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 278-0400  
TELEFAX: (212) 391-0525  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 457 amino acids

TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-328-500-9  
Query Match 39.7%; Score 1059; DB 4; Length 457;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVQLALPAPATQGGKGVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
DB 1 MNRGVPFRHLLVQLALPAPATQGGKGVLGKGGDTVELTCTASQKKSIOFHMKNNSQIK 60  
QY 61 ILGNQGSFLTQGPSKLNDRADSRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQL 120  
DB 61 ILGNQGSFLTQGPSKLNDRADSRSLWDQGNFPLIIKNLKIEDSDTYICEVEDQKEEVQL 120  
QY 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRKNIOGKKTLSVSOLELDDSG 180  
DB 121 LVFGLTANSDTHLLOQOSLTLTLESPPGSSPSVQCSPPRKNIOGKKTLSVSOLELDDSG 180  
QY 181 TWCTVQLQNKQKVEFKIDIVLAFQKAS 208  
DB 181 TWCTVQLQNKQKVEFKIDIVLAFQKAS 208  
RESULT 13  
US-08-466-368-4  
Sequence 4, Application US/08466368  
Patent No. 6093539  
GENERAL INFORMATION:  
APPLICANT: Maddon, Paul J.  
APPLICANT: Littman, Dan R.  
APPLICANT: Chess, Leonard  
APPLICANT: Axel, Richard  
APPLICANT: Weiss, Robin  
APPLICANT: McDougal, J. S.  
TITLE OF INVENTION: DNA ENCODING THE T CELL SURFACE PROTEIN  
TITLE OF INVENTION: T4 AND USE OF FRAGMENTS OF T4 IN THE TREATMENT OF AIDS  
NUMBER OF SEQUENCES: 21  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cooper & Dunham LLP  
STREET: 1185 Avenue of Americas  
CITY: New York  
STATE: New York  
COUNTRY: USA  
ZIP: 10036  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/466,368  
FILING DATE: 06-JUN-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: White, John P.  
REGISTRATION NUMBER: 28,678  
REFERENCE/DOCKET NUMBER: 24577-B1-B/JPM/AKC  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 212-391-0525  
TELEFAX: 212-278-0400  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 458 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-466-368-4  
Query Match 39.7%; Score 1059; DB 3; Length 458;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;

Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 NMRGVPFRLHLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIK 60  
Db 1 NMRGVPFRLHLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIK 60  
Qy 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLIIKNLKIEDSDTYICEVEDQKEEVL 120  
Db 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLIIKNLKIEDSDTYICEVEDQKEEVL 120  
Qy 121 LVFGLTANSDBTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDBTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Qy 181 TWTCTVLONQKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLONQKVEFKIDIVLAFQKAS 208

## RESULT 14

US-08-417-495-5  
Sequence 5, Application US/08417495  
Patent No. 5843728  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian et al.  
TITLE OF INVENTION: Redirection of Cellular Immunity by  
NUMBER OF SEQUENCES: 27  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Fish & Richardson  
STREET: 225 Franklin Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110-2804  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
COMPUTER: IBM PS/2 Model 502 or 55SX  
OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
SOFTWARE: Wordperfect (Version 5.0)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/417,495  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/203,866  
FILING DATE:  
APPLICATION NUMBER: US/07/847,566  
FILING DATE:  
APPLICATION NUMBER: 07/665,961  
FILING DATE: March 7, 1991  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Clark, Paul T.  
REGISTRATION NUMBER: 30,162  
REFERENCE/DOCKET NUMBER: 00786/119002  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (617) 542-5070  
TELEFAX: (617) 542-8906  
TELEX: 200154  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-417-495-5

Query Match 39.7%; Score 1059; DB 2; Length 462;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 NMRGVPFRLHLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIK 60

Db 1 NMRGVPFRLHLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIK 60  
Qy 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLIIKNLKIEDSDTYICEVEDQKEEVL 120  
Db 61 ILGNQGSFLTKGPSKLNDRADSRSLMDQGNFLIIKNLKIEDSDTYICEVEDQKEEVL 120  
Qy 121 LVFGLTANSDBTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Db 121 LVFGLTANSDBTHLQGSLLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
Qy 181 TWTCTVLONQKVEFKIDIVLAFQKAS 208  
Db 181 TWTCTVLONQKVEFKIDIVLAFQKAS 208

## RESULT 15

US-08-284-391B-5  
Sequence 5, Application US/08284391B  
Patent No. 5851828  
GENERAL INFORMATION:  
APPLICANT: Seed, Brian  
APPLICANT: Banapour, Babak  
APPLICANT: Romeo, Charles  
APPLICANT: Kolanus, Waldemar  
TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
NUMBER OF SEQUENCES: 53  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Clark & Ribling LLP  
STREET: 176 Federal Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02110  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/284,391B  
FILING DATE: 02-AUG-1994  
CLASSIFICATION: 514  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/195,395  
FILING DATE: 14-FEB-1994  
APPLICATION NUMBER: 07/847,566  
FILING DATE: 06-MAR-1992  
APPLICATION NUMBER: 07/665,961  
FILING DATE: 07-MAR-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Ribling, Karen L.  
REGISTRATION NUMBER: 35,238  
REFERENCE/DOCKET NUMBER: 00786/247001  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-428-0200  
TELEFAX: 617-428-7045  
TELEX:  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 462 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-284-391B-5

Query Match 39.7%; Score 1059; DB 2; Length 462;  
Best Local Similarity 99.5%; Pred. No. 2.5e-53;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 NMRGVPFRLHLVLTQALLPATQGNKVVLGKGGDTVELTCTASQKKSIOFHMKNSNOIK 60

Db 1 ||||| 60  
Qy 61 I|G|N|Q|S|F|L|T|K|P|S|K|N|D|R|A|D|S|R|R|S|L|W|D|Q|N|F|L|I|I|N|L|K|I|E|D|S|D|T|Y|I|C|E|V|E|D|Q|K|E|V|O|L 120  
Db 61 I|G|N|Q|S|F|L|T|K|P|S|K|N|D|R|A|D|S|R|R|S|L|W|D|Q|N|F|L|I|I|N|L|K|I|E|D|S|D|T|Y|I|C|E|V|E|D|Q|K|E|V|O|L 120  
Qy 121 L|V|F|G|L|T|A|N|S|D|T|H|L|L|Q|G|S|L|T|L|T|L|E|S|P|G|S|S|P|S|V|Q|C|R|S|P|R|G|K|N|I|Q|G|K|T|L|S|V|S|Q|L|E|L|Q|D|S|G 180  
Db 121 L|V|F|G|L|T|A|N|S|D|T|H|L|L|Q|G|S|L|T|L|T|L|E|S|P|G|S|S|P|S|V|Q|C|R|S|P|R|G|K|N|I|Q|G|K|T|L|S|V|S|Q|L|E|L|Q|D|S|G 180  
Qy 181 T|W|T|C|T|V|L|Q|N|Q|K|V|E|F|K|I|D|I|V|V|L|A|F|Q|K|A|S 208  
Db 181 T|W|T|C|T|V|L|Q|N|Q|K|V|E|F|K|I|D|I|V|V|L|A|F|Q|K|A|S 208

Search completed: November 21, 2004, 13:44:45  
Job time : 66.8175 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 ; Search time 222.482 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-3

Perfect score: 2668  
Sequence: 1 MNRGVPRHLLVQLALLP.....GTRLEIKLVPRSGHHHHH 508

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

Published Applications AA:\*

1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
18: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:\*  
19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Query length	ID	Description
1	1064	39.9	458	US-10-151-274-3	Sequence 3, Appli
2	1064	39.9	458	US-10-103-597A-39	Sequence 39, Appli
3	1064	39.9	458	US-10-188-444-39	Sequence 39, Appli
4	1064	39.9	458	US-10-207-655-170	Sequence 170, Appli
5	1062	39.8	458	US-08-681-219-27	Sequence 27, Appli
6	1062	39.8	458	US-09-230-111C-25	Sequence 25, Appli
7	1062	39.8	458	US-10-092-138-25	Sequence 25, Appli
8	1059	39.7	402	US-09-939-537-29	Sequence 29, Appli
9	1059	39.7	402	US-10-151-408-1	Sequence 1, Appli
10	1059	39.7	402	US-10-097-044A-1	Sequence 1, Appli
11	1059	39.7	402	US-10-769-247-1	Sequence 1, Appli
12	1059	39.7	457	US-09-891-119A-9	Sequence 9, Appli
13	1059	39.7	462	US-09-939-537-5	Sequence 5, Appli

14	1059	39.7	462	US-09-243-008-5	Sequence 5, Appli
15	1059	39.7	532	US-09-939-537-6	Sequence 6, Appli
16	1059	39.7	532	US-09-243-008-6	Sequence 6, Appli
17	1059	39.7	575	US-09-939-537-4	Sequence 4, Appli
18	1059	39.7	575	US-09-243-008-4	Sequence 4, Appli
19	1053	39.5	397	US-09-891-119A-2	Sequence 2, Appli
20	1051	39.4	530	US-08-485-163-5	Sequence 5, Appli
21	1051	39.4	530	US-08-766-995-4	Sequence 4, Appli
22	1050	39.4	432	US-08-485-163-3	Sequence 3, Appli
23	1050	39.4	432	US-09-766-995-2	Sequence 2, Appli
24	1048	39.3	310	US-08-485-163-7	Sequence 7, Appli
25	1048	39.3	310	US-09-766-995-6	Sequence 6, Appli
26	1035	38.8	203	US-09-939-537-31	Sequence 31, Appli
27	959.5	36.0	250	US-09-880-748-1952	Sequence 1952, Ap
28	959.5	36.0	250	US-10-293-418-1952	Sequence 1952, Ap
29	957.5	35.9	246	US-09-880-748-1268	Sequence 1268, Ap
30	957.5	35.9	246	US-10-293-418-1268	Sequence 1268, Ap
31	954	35.8	251	US-09-880-748-1238	Sequence 1238, Ap
32	954	35.8	251	US-10-293-418-1238	Sequence 1238, Ap
33	946	35.5	434	US-10-157-408-4	Sequence 4, Appli
34	946	35.5	434	US-10-097-044A-4	Sequence 4, Appli
35	946	35.5	434	US-10-769-247-4	Sequence 4, Appli
36	945.5	35.4	448	US-10-024-329-32	Sequence 32, Appli
37	940	35.2	370	US-09-759-841-6	Sequence 6, Appli
38	929	34.8	621	US-10-768-932A-4	Sequence 4, Appli
39	926	34.7	788	US-10-073-118-26	Sequence 26, Appli
40	923	34.6	590	US-09-934-060A-13	Sequence 13, Appli
41	923	34.6	720	US-09-934-060A-2	Sequence 2, Appli
42	923	34.6	720	US-09-934-060A-4	Sequence 4, Appli
43	917	34.4	453	US-10-768-932A-2	Sequence 2, Appli
44	917	34.4	453	US-10-768-932A-8	Sequence 8, Appli
45	916.5	34.4	476	US-10-768-932A-6	Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-10-151-274-3  
; Sequence 3, Application US/10151274  
; Publication No. US20030064071A1  
; GENERAL INFORMATION:  
; APPLICANT: Liltman, Dan R.  
; APPLICANT: Kwon, Douglas S.  
; APPLICANT: van Kooyk, Yvette  
; APPLICANT: Geljtenbeck, Theo  
; TITLE OF INVENTION: METHODS OF USING A FACILITATOR OF RETROVIRAL ENTRY  
; TITLE OF INVENTION: INTO  
; FILE REFERENCE: 1049-1-017  
; CURRENT APPLICATION NUMBER: US/10/151,274  
; CURRENT FILING DATE: 2002-05-20  
; PRIOR APPLICATION NUMBER: US/09/517,605  
; PRIOR FILING DATE: 2000-03-02  
; NUMBER OF SEQ ID NOS: 17  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 458  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-151-274-3

Query Match 39.9%; Score 1064; DB 14; Length 458;  
Best Local Similarity 100.0%; Pred. No. 7.9e-51;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALLPAAATGKKVYVETCTTASOKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRHLLVQLALLPAAATGKKVYVETCTTASOKSIOFHMKNSNOIK 60  
QY 61 ILNGSGSLTGTGPKSLNDRASRSLSMDQGNFLLIKLKLTEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGSGSLTGTGPKSLNDRASRSLSMDQGNFLLIKLKLTEDSDTYICEVEDQKEVQL 120

```

0y      121  LVFGLTANSDFHLLQGOSLTLTLESPPGSSPSVQCSPPGKXNOGGKTLVSQLELODSG 180
Db      121  LVFGLTANSDFHLLQGOSLTLTLESPPGSSPSVQCSPPGKXNOGGKTLVSQLELODSG 180

0y      181  TWCTCTVLONQKKVEFKIDIVLAFOKAS 208
Db      181  TWCTCTVLONQKKVEFKIDIVLAFOKAS 208

RESULT 2
US-10-103-597A-39
; Sequence 39, Application US/10103597A
; Publication NO. US20030096432A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142
; CURRENT APPLICATION NUMBER: US/10/103,597A
; CURRENT FILING DATE: 2002-10-17
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 992352.1
; PRIOR FILING DATE: 1998-09-21
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 39
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-103-597A-39

```

```

Query Match Similarity      39.9%: Score 1064; DB 14; Length 458;
Best Local Similarity      100.0%: Pred. No. 7,9e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0

OY      1  MNRGVPFRHLILVLQIALLPAAATGCKVVLGKKGDIVBELCTASQKKSIOFHMKNSNOIK 60
DB      1  MNRGVPFRHLILVLQIALLPAAATGCKVVLGKKGDIVBELCTASQKKSIOFHMKNSNOIK 60
OY      61  ILNGSGFLTKPSPSKINDPADRRSLMDGNFPPLIKNLKIESDPYICGVEPQKEEVOL 120
DB      61  ILNGSGFLTKPSPSKINDPADRRSLMDGNFPPLIKNLKIESDPYICGVEPQKEEVOL 120
OY      121  LVFGITANDPDLHLOGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
DB      121  LVFGITANDPDLHLOGQSLTTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180
OY      181  TWTCTVLQNKQKVEFKIDIVLAFQKAS 208
DB      181  TWTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 3
US-10-188-444-39
; Sequence 39, Application US/10188444
; Publication No. US20030104635A1
; GENERAL INFORMATION:
; APPLICANT: Jakobsen, Bent Karsten
; TITLE OF INVENTION: Screening Methods
; FILE REFERENCE: 102286.142 (CIP)
; CURRENT APPLICATION NUMBER: US/10/188,444
; CURRENT FILING DATE: 2002-07-02
; PRIOR APPLICATION NUMBER: PCT/GB00/03579
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: GB 9922352.1
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: FastSeq for Windows Version 4.0
; LENGTH: 458
; TYPE: prt
; ORGANISM: Homo sapiens

```

[illegible]

```

RESULT 4
US-10-207-655-170
; Sequence 170, Application US/10207655
; Publication No. US20030118592A1
; GENERAL INFORMATION:
; APPLICANT: Ledbetter, Jeffrey A.
; APPLICANT: Hayden-Ledbetter, Martha S.
; TITLE OF INVENTION: BINDING DOMAIN-IMMUNOGLOBULIN FUSION PROTEINS
; FILE REFERENCE: 390069.401C1
; CURRENT APPLICATION NUMBER: US/10/207,655
; CURRENT FILING DATE: 2002-07-25
; NUMBER OF SEQ ID NOS: 426
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 170
; LENGTH: 458
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-207-655-170

Query Match          39.9%; Score 1064; DB 14; Length 458;
Best Local Similarity 100.0%; Pred. No. 7, 9e-51;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1  MNRGPPRHLLIVQLALLPATGGKRVYLGKKDYTELTCTASQKSIQPHMKNQIQK 60
DB      1  MNRGPPRHLLIVQLALLPATGGKRVYLGKKDYTELTCTASQKSIQPHMKNQIQK 60
QY      61  ILNGGSEFLTGPKSKMDNRADSRSLMDQGNFLIKNLKIEDPTYICEVEDQKEEYQL 120
DB      61  ILNGGSEFLTGPKSKMDNRADSRSLMDQGNFLIKNLKIEDSDTYICEVEDQKEEYQL 120
QY      121  LVFGITANSPTHLIQGSLTTLTLESPGSSPSVQCRSPKNGNIQGGKTLVSQLELDSG 180
DB      121  LVFGITANSPTHLIQGSLTTLTLESPGSSPSVQCRSPKNGNIQGGKTLVSQLELDSG 180
QY      181  TWCTCTVLQNKQKVEFKIDIVLAQKAS 208
DB      181  TWCTCTVLQNKQKVEFKIDIVLAQKAS 208

RESULT 5
US-08-681-219-27
; Sequence 27, Application US/08681219
; Publication No. US20020058607A1
; GENERAL INFORMATION:
; APPLICANT: Takaaki Sato and Junn Yanagisawa
; TITLE OF INVENTION: COMPOUNDS THAT INHIBIT THE INTERACTION BETWEEN
; TITLE OF INVENTION: SIGNAL-TRANSDUCING PROTEINS AND THE GAGF
; TITLE OF INVENTION: (P2D/DHR) DOMAIN AND USES THEREOF
; NUMBER OF SEQUENCES: 35

```



```

;
; NUMBER OF SEQUENCES: 53
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Clark & Elbing LLP
; STREET: 176 Federal Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/939,537
; FILING DATE: 24-Aug-2001
; CLASSIFICATION: <Unknown>
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/284,391
; FILING DATE: 02-AUG-1994
; APPLICATION NUMBER: 08/195,395
; FILING DATE: 14-FEB-1994
; APPLICATION NUMBER: 07/847,566
; FILING DATE: 06-MAR-1992
; APPLICATION NUMBER: 07/665,961
; FILING DATE: 07-MAR-1991
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Elbing, Karen L
; REGISTRATION NUMBER: 35,238
; REFERENCE/DOCKET NUMBER: 00786/247001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-428-0200
; TELEFAX: 617-428-7045
;
; TELEEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 29:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 398 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; SEQUENCE DESCRIPTION: SEQ ID NO: 29:
US-09-939-537-29

Query Match          39.7%; Score 1059; DB 10; Length 398;
Best Local Similarity 99.5%; Pred. No. 1.3e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAAATQGGKGVVLTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAAATQGGKGVVLTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVO 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVO 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRRPRGKNIQGGKTLTSLVSOLELDSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRRPRGKNIQGGKTLTSLVSOLELDSG 180
QY 181 TWCTCTVONOKKVEFKIDIVVLAFOKAS 208
DB 181 TWCTCTVONOKKVEFKIDIVVLAFOKAS 208

RESULT 9
US-10-157-408-1
; Sequence 1, Application US/10157408
; Publication No. US20030104535A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
; Gregory, Timothy J.
; TITLE OF INVENTION: Adhesion Variants
```

```

;
; NUMBER OF SEQUENCES: 25
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genentech, Inc.
; STREET: 460 Point San Bruno Blvd
; CITY: South San Francisco
; STATE: California
; COUNTRY: USA
; ZIP: 94080
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: patin (Genentech)
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/157,408
; FILING DATE: 28-May-2002
; CLASSIFICATION: 435
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/457,918
; FILING DATE: 1-JUN-1995
; APPLICATION NUMBER: 08/236311
; FILING DATE: 02-MAY-1994
; APPLICATION NUMBER: 07/936190
; FILING DATE: 26-AUG-1992
; APPLICATION NUMBER: 07/842777
; FILING DATE: 18-FEB-1992
; APPLICATION NUMBER: 07/250785
; FILING DATE: 28-SEP-1988
; APPLICATION NUMBER: 07/104329
; FILING DATE: 02-OCT-1987
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Kubinec, Jeffrey S.
; REGISTRATION NUMBER: 36,575
; REFERENCE/DOCKET NUMBER: P0444P1C3
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415/225-8228
; TELEFAX: 415/952-9881
; TELEEX: 910/371-7168
;
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-10-157-408-1

Query Match          39.7%; Score 1059; DB 14; Length 402;
Best Local Similarity 99.5%; Pred. No. 1.3e-50;
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPRHLLVQLALPAAATQGGKGVVLTCTASQKKSIOFHMKNNOIK 60
DB 1 MNRGVPRHLLVQLALPAAATQGGKGVVLTCTASQKKSIOFHMKNNOIK 60
QY 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVO 120
DB 61 ILGNQGSFLTKGPSKLNDRADSRSLWDQGNFPLIINKLKIEDSDTYICEVEDQKEVO 120
QY 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRRPRGKNIQGGKTLTSLVSOLELDSG 180
DB 121 LVFGLTANSDTHLLQGQSLTLTLSPGSSPSVOCRRPRGKNIQGGKTLTSLVSOLELDSG 180
QY 181 TWCTCTVONOKKVEFKIDIVVLAFOKAS 208
DB 181 TWCTCTVONOKKVEFKIDIVVLAFOKAS 208

RESULT 10
US-10-097-044A-1
; Sequence 1, Application US/10097044A
; Publication No. US20030143320A1
; GENERAL INFORMATION:
; APPLICANT: Capon, Daniel J.
```

Gregory, Timothy J.  
TITLE OF INVENTION: Adheseon Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/097,044A  
FILING DATE: 28-May-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936130  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-097-044A-1

Query Match 39.7%; Score 1059; DB 14; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1.3e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGVVLAGKGDVVELTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATQGVVLAGKGDVVELTCTASQKKSIOFHMKNNOIK 60  
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWCTCTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWCTCTVLQNKQKVEFKIDIVVLAFOKAS 208

RESULT 11  
US-10-769-247-1  
; Sequence 1, Application US/10769247  
; Publication No. US20040197809A1

GENERAL INFORMATION:  
APPLICANT: Capon, Daniel J.  
Gregory, Timothy J.  
TITLE OF INVENTION: Adheseon Variants  
NUMBER OF SEQUENCES: 25  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 5.25 inch, 360 Kb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: patin (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/769,247  
FILING DATE: 30-Jan-2004  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/457,918  
FILING DATE: 1-JUN-1995  
APPLICATION NUMBER: 08/236311  
FILING DATE: 02-MAY-1994  
APPLICATION NUMBER: 07/936190  
FILING DATE: 26-AUG-1992  
APPLICATION NUMBER: 07/842777  
FILING DATE: 18-FEB-1992  
APPLICATION NUMBER: 07/250785  
FILING DATE: 28-SEP-1988  
APPLICATION NUMBER: 07/104329  
FILING DATE: 02-OCT-1987  
ATTORNEY/AGENT INFORMATION:  
NAME: Kubinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0444P1C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 402 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
US-10-769-247-1

Query Match 39.7%; Score 1059; DB 17; Length 402;  
Best Local Similarity 99.5%; Pred. No. 1.3e-50;  
Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLLVQLALLPAATQGVVLAGKGDVVELTCTASQKKSIOFHMKNNOIK 60  
DB 1 MNRGVPFRHLLVQLALLPAATQGVVLAGKGDVVELTCTASQKKSIOFHMKNNOIK 60  
QY 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120  
DB 61 ILNGQSFLLTGKPSKLNDRADSRSLMDQGNFPLIIKNLKIEDSDTYICEVEDQKEEYQL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLSPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWCTCTVLQNKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWCTCTVLQNKQKVEFKIDIVVLAFOKAS 208

RESULT 12  
US-09-891-119A-9

Sequence 9, Application US/09891119A  
 Publication No. US20040013683A1  
 GENERAL INFORMATION:  
 APPLICANT: Maddon, Paul J.  
 TITLE OF INVENTION: DERIVATIVES OF SOLUBLE T-4  
 FILE REFERENCE: 24577-CY-B  
 CURRENT APPLICATION NUMBER: US/09/891.119A  
 CURRENT FILING DATE: 2001-06-25  
 NUMBER OF SEQ ID NOS: 22  
 SOFTWARE: PatentIn version 3.1  
 SEQ ID NO 9  
 LENGTH: 457  
 TYPE: PRT  
 ORGANISM: human  
 US-09-891-119A-9

Query Match 39.7%; Score 1059; DB 11; Length 457;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQKRVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLVLVQLALLPAATQKRVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60  
 QY 61 ILNQGSPFLTKGSPSKLNDRAISRSLWDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNQGSPFLTKGSPSKLNDRAISRSLWDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWCTVQLNQKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWCTVQLNQKQKVEFKIDIVLAFQKAS 208

## RESULT 13

US-09-939-537-5  
 Sequence 5, Application US/09939537  
 Publication No. US20030138410A1  
 GENERAL INFORMATION:

APPLICANT: Seed, Brian  
 Banapour, Babak  
 Romeo, Charles  
 Kolanus, Waldemar  
 TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INJECTED  
 CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
 NUMBER OF SEQUENCES: 53  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Clark & Elbling LLP  
 STREET: 176 Federal Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/939.537  
 FILING DATE: 24-Aug-2001  
 CLASSIFICATION: <Unknown>  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: 08/284.391  
 FILING DATE: 02-AUG-1994  
 APPLICATION NUMBER: 08/195.395  
 FILING DATE: 14-FEB-1994  
 APPLICATION NUMBER: 07/847.566  
 FILING DATE: 06-MAR-1992  
 APPLICATION NUMBER: 07/665.961

FILING DATE: 07-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbling, Karen L.  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/247001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-428-0200  
 TELEFAX: 617-428-7045  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 462 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 US-09-939-537-5

Query Match 39.7%; Score 1059; DB 10; Length 462;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MNRGVPFRHLVLVQLALLPAATQKRVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60  
 DB 1 MNRGVPFRHLVLVQLALLPAATQKRVLGKGGDTVELTCTASQKKSIOFHMKNNOIK 60  
 QY 61 ILNQGSPFLTKGSPSKLNDRAISRSLWDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNQGSPFLTKGSPSKLNDRAISRSLWDQGNFPLIKNLKIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLOQOSLTLTLESPGSSPSVQCSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWCTVQLNQKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWCTVQLNQKQKVEFKIDIVLAFQKAS 208

## RESULT 14

US-09-243-008-5  
 Sequence 5, Application US/09243008  
 Publication No. US20040005334A1  
 GENERAL INFORMATION:

APPLICANT: Seed, Brian et al.  
 TITLE OF INVENTION: Redirection of Cellular Immunity by  
 Receptor Chimeras  
 NUMBER OF SEQUENCES: 40  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Fish & Richardson P.C.  
 STREET: 225 Franklin Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02110-2804  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: 3.5" Diskette, 1.44 Mb  
 COMPUTER: IBM PS/2 Model 50Z or 55SX  
 OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)  
 SOFTWARE: Wordperfect (Version 5.0)  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/243.008  
 FILING DATE: 02-Feb-1999  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/394.176  
 FILING DATE: SEPTEMBER 11, 1995  
 APPLICATION NUMBER: 08/203.866  
 FILING DATE: February 28, 1994  
 APPLICATION NUMBER: 07/847.566  
 FILING DATE: March 6, 1992  
 APPLICATION NUMBER: 07/665.961  
 FILING DATE: March 7, 1991

ATTORNEY/AGENT INFORMATION:  
 NAME: Karen F. Lech, Ph.D  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/270001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (617) 542-5070  
 TELEFAX: (617) 542-8906  
 TELEX: 200154  
 INFORMATION FOR SEQ ID NO: 5:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 462 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 5:  
 US-09-243-008-5

Query Match 39.7%; Score 1059; DB 11; Length 462;  
 Best Local Similarity 99.5%; Pred. No. 1.5e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPFRLILVQLALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNOIK 60  
 DB 1 MRGVPFRLILVQLALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNOIK 60  
 QY 61 ILNGGSPFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEYVL 120  
 DB 61 ILNGGSPFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEYVL 120  
 QY 121 LVFGLTANSDFHLQGGSLITLLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLQGGSLITLLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONOKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLONOKKVEFKIDIVLAFOKAS 208

RESULT 15  
 US-09-939-537-6  
 ; Sequence 6, Application US/09939537  
 ; Publication No. US20030138410A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Seed, Brian  
 ; Banapour, Babak  
 ; Kolanus, Charles  
 ; TITLE OF INVENTION: TARGETED CYTOLYSIS OF HIV-INFECTED  
 ; CELLS BY CHIMERIC CD4 RECEPTOR-BEARING CELLS  
 ; NUMBER OF SEQUENCES: 53  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSSEE: Clark & Elbing LLP  
 ; STREET: 176 Federal Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02110  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM Compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FastSeq for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/939,537  
 ; FILING DATE: 24-Aug-2001  
 ; CLASSIFICATION: <Unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 08/284,391  
 ; FILING DATE: 02-AUG-1994  
 ; APPLICATION NUMBER: 08/195,395  
 ; FILING DATE: 14-FEB-1994  
 ; APPLICATION NUMBER: 07/847,566  
 ; FILING DATE: 06-MAR-1992

APPLICATION NUMBER: 07/665,961  
 FILING DATE: 07-MAR-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elbing, Karen L  
 REGISTRATION NUMBER: 35,238  
 REFERENCE/DOCKET NUMBER: 00786/247001  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-428-0200  
 TELEFAX: 617-428-7045  
 TELEX: <Unknown>  
 INFORMATION FOR SEQ ID NO: 6:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 532 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 SEQUENCE DESCRIPTION: SEQ ID NO: 6:  
 US-09-939-537-6

Query Match 39.7%; Score 1059; DB 10; Length 532;  
 Best Local Similarity 99.5%; Pred. No. 1.7e-50;  
 Matches 207; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MRGVPFRLILVQLALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNOIK 60  
 DB 1 MRGVPFRLILVQLALIPAAATQGNKVLGKGGDTVELTCTASQKKSIOFHKNSNOIK 60  
 QY 61 ILNGGSPFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEYVL 120  
 DB 61 ILNGGSPFLTKGSPSKLNDRAISRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEEYVL 120  
 QY 121 LVFGLTANSDFHLQGGSLITLLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDFHLQGGSLITLLESPPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLONOKKVEFKIDIVLAFOKAS 208  
 DB 181 TWTCTVLONOKKVEFKIDIVLAFOKAS 208

Search completed: November 21, 2004, 14:03:56  
 Job time : 227.482 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:25:21 : Search time 2.58212 Seconds  
(without alignments)  
694.642 Million cell updates/sec

Title: US-09-936-702-1

Perfect score: 28  
Sequence: 1 GCGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues  
Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_23Sep04:\*  
1: geneeqp1980s:\*  
2: geneeqp1990s:\*  
3: geneeqp2000s:\*  
4: geneeqp2001s:\*  
5: geneeqp2002s:\*  
6: geneeqp2003as:\*  
7: geneeqp2003bs:\*  
8: geneeqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	5	2	AAR72707 linker fo
2	28	100.0	5	2	AAR34034 Linking s
3	28	100.0	5	2	AAR95082 scfv spac
4	28	100.0	5	2	AAW17094 Gly(4)-Se
5	28	100.0	5	2	AAW19543 Chimeric
6	28	100.0	5	2	AAV02127 Peptide 1
7	28	100.0	5	2	AAV25357 IFNAR2/IF
8	28	100.0	5	2	AAV33587 VH-VL dom
9	28	100.0	5	2	AAV43496 linker fo
10	28	100.0	5	3	AAW83210 Peptide 1
11	28	100.0	5	3	AAW6226 Expressio
12	28	100.0	5	3	AAV4917 linker fr
13	28	100.0	5	3	AAV43750 linker us
14	28	100.0	5	3	AAW14535 Peptide 1
15	28	100.0	5	3	AAW00156 linker us
16	28	100.0	5	4	AAW06269 Glycine 1
17	28	100.0	5	4	AAE11873 Spacer pe
18	28	100.0	5	4	AAW57374 Flexible
19	28	100.0	5	4	ABW56480 Human sin
20	28	100.0	5	4	AAW62343 linker pe
21	28	100.0	5	4	AAW50848 Fluoresce
22	28	100.0	5	4	AAW57784 Peptide 1
23	28	100.0	5	4	AAW55527 Peptide 1
24	28	100.0	5	4	AAW81033 linker pe
25	28	100.0	5	4	AAW62613 linker pe

26	28	100.0	5	4	AAU05173 Pain-rel
27	28	100.0	5	4	AAU05172 Pain-rel
28	28	100.0	5	4	AAW63011 Peptide 1
29	28	100.0	5	4	AAW68571 Linker pe
30	28	100.0	5	4	AAW74580 Context-d
31	28	100.0	5	4	AAW08730 Peptide 1
32	28	100.0	5	5	ABG31871 Spacer pe
33	28	100.0	5	5	AAU76043 Synthetic
34	28	100.0	5	5	ABW79971 Linker pe
35	28	100.0	5	5	ABW79970 Linker pe
36	28	100.0	5	5	ABW52872 Flexible
37	28	100.0	5	5	ABG31035 Binding d
38	28	100.0	5	5	ABJ05006 VH and VL
39	28	100.0	5	5	ABW48193 Peptide 1
40	28	100.0	5	5	ABW60659 Polyclonal
41	28	100.0	5	5	AAW14387 Linker pe
42	28	100.0	5	5	AAE19822 Linker pe
43	28	100.0	5	5	AAE25959 Linker pe
44	28	100.0	5	5	ABJ03945 Peptide 1
45	28	100.0	5	5	AAU75495 Human Gly

## ALIGNMENTS

RESULT 1  
ID AAR72707 standard; peptide; 5 AA.

XX AAR72707;

DT 31-OCT-1995 (first entry)

DE Linker for apo A-I and apo B-100 fusion polypeptide.

KW Apo A-I, LDL cholesterol; low density lipoprotein; fusion polypeptide; linker.

XX Synthetic.

XX US408038-A.

PN 18-APR-1995.

XX 08-OCT-1992; 92US-00959946.

PR 09-OCT-1991; 91US-00774633.

XX 18-JUN-1992; 92US-00901706.

PA (SCRI) SCRIPPS RES INST.

PI Witczum JL, Koduri KR, Young SG, Smith RS, Curtiss LK;

DR WPI; 1993-134378/16.

XX Polypeptide mimic of native apo B-100 and native apo A-I - useful in assays for LDL and HDL in plasma samples.

PT Disclosure; Col 18; 41pp; English.

PS A dispersible apo A-I/B-100 fusion polypeptide is claimed which contains a first AA sequence of apo A-I (see AAR72605) and that includes at least AA sequence positions 120-135 (see AAR72606). The two sequences are operatively linked. An exemplary linking sequence is AAR72707 whose encoding DNA can be ligated between an apo A-I and a B-100 encoding DNA sequence

CC XX

CC XX

SO

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

# RESULT 2

AAR34034  
 ID AAR34034 standard; protein; 5 AA.  
 XX  
 AC AAR34034;

XX  
 DT 25-MAR-2003 (revised)  
 DT 13-AUG-1993 (first entry)

XX  
 DE Linking sequence whose encoding DNA can be ligated between an apo A-I-  
 and a B-100-encoding DNA sequence.

XX  
 KW Lipoprotein; apoprotein; B-100; A-I; LDL; HDL; assay.

XX  
 OS Synthetic.

XX  
 PN WO9307165-A1.

XX  
 PD 15-APR-1993.

XX  
 PF 09-OCT-1992; 92WO-US008634.

XX  
 PR 09-OCT-1991; 91US-00774633.

XX  
 PR 18-JUN-1992; 92US-00901706.

XX  
 PR 08-OCT-1992; 92US-00959946.

XX  
 PA (SCRI ) SCRIPPS RES INST.

XX  
 PI Smith RS, Curtles LK, Koduri KR, Witzum JL, Young SG;

XX  
 DR WPI; 1993-134378/16.

XX  
 PT Polypeptide mimic of native apo B-100 and native apo A-I - useful in  
 assay for LDL and HDL in plasma samples.

XX  
 PS Disclosure; Page 14 and page 35; 137pp; English.

XX  
 CC The inventors claim a portion of the polypeptide contg. apo B-100 that  
 immunoreacts with antibodies secreted by the hybridoma MB47 having ATCC

XX  
 CC Accession No. 8746. Polypeptides specifically claimed include residues  
 CC 217-297, 216-310, 216-331, 216-352, 216-377, 1-377, 205-297, 173-297, 140

XX  
 CC -297. DNA sequences encoding the polypeptides are also claimed. Also  
 CC claimed are a fusion polypeptide that contains: (a) a first amino

XX  
 CC acid residue sequence up to 250 residues in length that includes residues  
 CC 120-135 of apo A-I, (b) a second amino acid residue sequence up to 375

XX  
 CC residues in length that includes residues 217-297 of apo B-100 and DNA  
 CC encoding it. (Updated on 25-MAR-2003 to correct PN field.) (Updated on 25

XX  
 CC -MAR-2003 to correct PR field.)

XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

RESULT 3  
 AAR95062  
 ID AAR95062 standard; peptide; 5 AA.  
 XX  
 AC AAR95062;

XX  
 DT 18-AUG-1996 (first entry)  
 XX

DE scFv spacer peptide.  
 XX  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer; scFv;  
 KW single chain antibody.

XX  
 OS Synthetic.

XX  
 PN WO9613599-A1.

XX  
 PD 09-MAY-1996.

XX  
 PF 31-OCT-1995; 95WO-EP004270.

XX  
 PR 01-NOV-1994; 94EP-00810627.

XX  
 PA (WELS/) WELS W.

XX  
 PI Wels W, Fominaya J;

XX  
 DR WPI; 1996-239505/24.

XX  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 cell.

XX  
 PS Disclosure; Page 8; 106pp; English.

XX  
 CC A flexible spacer peptide (AAR95062) is used to link the light chain  
 CC variable domain to the heavy chain variable domain of a single chain

XX  
 CC recombinant antibody (scFv). The scFv may be derived from a monoclonal  
 CC antibody, e.g. Mab FRP5, and forms the ligand domain of a multidomain

XX  
 CC protein (see also AAR95053 and AAR95056-58) that is used with an effector  
 CC nucleic acid in a novel nucleic acid transfer system suitable for gene

XX  
 CC therapy. The ligand domain has a target cell recognition function and  
 CC allows cellular internalization of the multidomain protein/nucleic acid

XX  
 CC complex

XX  
 SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

RESULT 4  
 AAM17094  
 ID AAM17094 standard; peptide; 5 AA.  
 XX  
 AC AAM17094;

XX  
 DT 14-SEP-1999 (first entry)  
 DT  
 XX  
 DE Gly(4)-Ser linker peptide for chimeric protein construct.

XX  
 KW Haematopoietic protein; human; granulocyte-colony stimulating factor;  
 KW G-CSF; interleukin; c-mpl ligand; linker; gene therapy; aplastic anaemia;  
 KW stem cell expansion; leukopaenia; neutropaenia; vector; bone marrow;  
 KW thrombocytopenia; blood cell activation; growth.

XX  
 OS Synthetic.

XX  
 PN WO9712985-A2.

XX  
 PD 10-APR-1997.

XX  
 PF 04-OCT-1996; 96WO-US015774.

XX  
 PR 05-OCT-1995; 95US-0004834P.

(SEAR) SEARLE & CO G. D.

Peng Y, Staten NR, Baum CM, Summers NL, Caparon MH, Bauer SC, Zurfluh L, McKeam JP, Klein BK, Lee SC, McWhorter CA, Giri JG, WPI, 1997-226228/20.

Multi-functional haematopoietic receptor agonists - used to stimulate the production of haematopoietic cells in patients.

Disclosure; Page 33; 616pp; English.

The invention relates to a novel haematopoietic protein (HP) comprising an amino acid (AA) sequence of Formula: R1-L1-R2; R2-L1-R1; R1-R2; or R2-R1, where R1 and R2 are independently selected from: (i) a modified human granulocyte-colony stimulating factor (hG-CSF) AA sequence; (ii) a modified human interleukin-3 (hIL-3) AA sequence; (iii) a modified human c-mpl ligand; and a colony stimulating factor (CSF); and L1 = a linker capable of linking R1 to R2. This sequence represents an example of a linker used to construct the proteins of the invention. Vectors comprising the nucleic acid molecules are useful for the recombinant production of HP. The nucleic acid molecules are useful in gene therapy. The HPs are useful for stimulating the production of haematopoietic cells in patients, selective ex vivo expansion of stem cells and for treatment of haematopoietic disorders. Disorders that can be treated include leukopaemia, neutropaenia, aplastic anaemia and thrombocytopaenia. In vitro uses include the ability to stimulate bone marrow and blood cell activation and growth before infusion into the patients

Sequence 5 AA;

Query Match	100.0%	Score 28;	DB 2;	Length 5;
Best Local Similarity	100.0%	Pred. NO. 1.7e+06;		
Matches	5;	Conservative 0;	Indels 0;	Gaps 0

QY 1 GGGS 5  
1 1 1 1 1 1  
Db 1 GGGS 5

```

RESULT 5
AAW19543
ID AAW19543 standard; peptide; 5 AA.

```

AC	AAW19543;
XX	
DT	19-FEB-1998 (first entry)
...	

DE Chimeric protein pentapeptide linker for the MBP moiety and PE moiety.

KM Pseudomonas exotoxin; myelin basic protein; chimeric protein;  
KM autoimmune disease; multiple sclerosis; human.  
...

OS Synthetic.

PN WO9719179-A1.

PD 29-MAY-1997

PF 17-NOV-1996; 96WO-IL000151.

PR 17-NOV-1995; 95IL-00116044.

XX  
XX  
XXXXX ) XXXXXXXX  
XXXXX ) XXXXXXXX

**XX**

PI Yarkoni S;  
yy

DR WPI; 1997-298116/27.  
YY

PT New Pseudomonas exotoxin-myeelin basic protein chimeric proteins - used  
PT for the treatment of auto:immune diseases, particularly multiple  
PT sclerosis.

PS Claim 6; Page 29; 54pp; English.

CC New chimeric proteins have been developed comprising a Pseudomonas  
CC aeruginosa exotoxin (PE) moiety linked to a myelin basic protein (MBP)  
CC moiety selected from: (a) MBP; (b) amino acids 69-88 of guinea-pig MBP or  
CC an antigenic portion; (c) amino acids 84-102 of human MBP or an antigenic  
CC portion; (d) amino acids 143-168 of human MBP or an antigenic portion;  
CC and (e) an amino acid sequence in which one or more amino acids have been  
CC deleted, added, substituted or mutated in the amino acid sequences of  
CC (a), (b), (c), or (d), the modified sequences retaining at least 75%  
CC homology with the amino acid sequences. The present sequence represents  
CC the preferred pentapeptide linker used to link the MBP moiety and PE  
CC moiety in a chimeric protein. The chimeric proteins can be used for the  
CC treatment of autoimmune diseases such as multiple sclerosis. The chimeric  
CC proteins can specifically target and kill MBP specific T cells while  
CC having no effect on non-target cells

SQ Sequence 5 AA;

Query Match	100.0%	Score 28;	DB 2;	Length 5;
Best Local Similarity	100.0%	Pred. NO.17e+06;		
Matches 5;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0

QY	1	GGGS	5
Db	1	GGGS	5

```

RESULT 6
AAY02127
ID      AAY02127  standard; protein; 5 AA.
....

```

AC AAY02127;

DT 16-JUL-1999 (first entry)

Peptide linker used to make multifunctional proteins.

KM Angiostatin; endostatin; interferon; thrombospondin;  
KM Interferon-inducible protein; platelet factor 4; anti-angiogenic;  
KM anti-tumor; multifunctional protein; angiogenic-mediated disease; cancer  
KM diabetic retinopathy; macular degeneration; arthritis;  
KM tumor cell production; peptide linker.

OS Homo sapiens.

PN WO9916889-A1.

PD 08-APR-1999

PF 30-SEP-1998; 98WO-US020464.

PR 01-OCT-1997; 97US-0060609P.  
yy

PA (SEAR ) SEARLE & CO G D.  
VV

PI Bolanowski MA, Caparon MH, Casperson GF, Gregory SA, Klein BK;  
DT Moberg TD.

XX  
WPB: 1999-355099/21

PT New multifunctional proteins useful for treating angiogenic-mediated diseases.

PS Disclosure; Page 111; 121pp; English.

CC The specification describes multifunctional proteins which comprise

CC interferon-inducible protein and platelet factor 4, and have anti-

CC angiogenic and/or anti-tumor activity. The multifunctional protein may  
 CC exhibit useful properties such as having similar or greater biological  
 CC activity when compared to a single factor or by having improved half-life  
 CC or decreased adverse side effects, or a combination of these properties.  
 CC The proteins can be used for treating an angiogenic-mediated disease,  
 CC e.g., cancer, diabetic retinopathy, macular degeneration, or arthritis.  
 CC (They can also be used for inhibiting the production of tumor cells  
 CC characteristic of lung, breast, ovarian, prostate, pancreatic, gastric,  
 CC colon, renal, bladder cancers; melanoma, hepatoma, sarcoma and lymphoma)  
 CC in a patient and for inhibiting tumor growth. AAY02125-32 represent  
 CC peptide linkers used to make the multifunctional proteins of the  
 CC invention

XX SQ Sequence 5 AA;

XX QY 1 GGGGS 5  
 XX 1 GGGGS 5

XX DB 1 GGGGS 5

XX RESULT 7  
 XX AAY25357  
 XX ID AAY25357 standard; peptide; 5 AA.  
 XX XX  
 XX AAY25357;  
 XX XX  
 XX DT 06-SEP-1999 (first entry)  
 XX XX  
 XX DE IFNAR2/IFN-beta complex peptide fragment 1.  
 XX XX  
 XX KM IFNAR2; IFN-beta; type I interferon; IFNAR/IFN complex; IFN, antiviral;  
 XX KM human interferon alpha/beta receptor; anticancer; immunomodulatory;  
 XX KM anti-arthritis; antidiabetic; treatment; hepatitis; viral infection;  
 XX KM hairy cell leukemia; Kaposi's sarcoma; multiple myeloma; cancer; lupus;  
 XX KM diabetes; multiple sclerosis; rheumatoid arthritis; myasthenia gravis;  
 XX KM acquired immune deficiency syndrome.  
 XX XX  
 XX OS Synthetic.  
 XX XX  
 XX PN WO9932141-A1.  
 XX XX  
 XX PD 01-JUL-1999.  
 XX XX  
 XX PF 18-DEC-1998; 98WO-US026926.  
 XX XX  
 XX PR 19-DEC-1997; 97US-0068295P.  
 XX XX  
 XX PA (ISTF) ARS APPLIED RES SYSTEMS HOLDING NV.  
 XX PA (MCIN/) MCINNIS P G.  
 XX XX  
 XX PI Tepper M, Cunningham M, Sherris D, El Tayar N, McKenna S;  
 XX DR WPI; 1999-405115/34.  
 XX XX  
 XX PT Prolonging in vivo activity of type I interferon by complexing.  
 XX XX  
 XX PS Example 8; Page 76; 86pp; English.  
 XX XX  
 CC This invention describes a novel method for prolonging the in vivo effect  
 CC of type I interferon (IFN) by administering IFN as a complex (A) with a  
 CC subunit (I) of the human interferon alpha/beta receptor (IFNAR). The  
 CC product of the invention has antiviral, anticancer, immunomodulatory,  
 CC anti-arthritis and antidiabetic activity. (A) have the antiviral,  
 CC anticancer and immunomodulating activities of IFN, e.g., for treating  
 CC hepatitis and other viral infections, hairy cell leukemia, Kaposi's  
 CC sarcoma, multiple myeloma and other cancers, multiple sclerosis,  
 CC rheumatoid arthritis, myasthenia gravis, diabetes, acquired immune  
 CC deficiency syndrome and lupus. When complexed in (A), the storage life of  
 CC IFN is increased (i.e. it is stabilized against oligomerization, without

CC the need for storage at acidic pH) and its biological effect is  
 CC potentiated

XX SQ Sequence 5 AA;

XX QY 1 GGGGS 5  
 XX 1 GGGGS 5

XX DB 1 GGGGS 5

XX RESULT 8  
 XX AAY33597  
 XX ID AAY33597 standard; protein; 5 AA.  
 XX XX  
 XX AC AAY33597;  
 XX XX  
 XX DT 20-DEC-1999 (first entry)  
 XX XX  
 XX DE VH-VL domain linker peptide #9.  
 XX XX  
 XX KM Antigen binding; single chain; variable domain; VH domain; light chain;  
 XX KM heavy immunoglobulin chain; VL domain; anticancer; antiviral; tumor;  
 XX KM antibacterial; antimalarial; antiinflammatory; treatment; prevention;  
 XX KM diagnosis; vaccine; autoimmune disease; inflammation; blood disorder;  
 XX KM transplant rejection; arthritis; nervous system disorder; infection.  
 XX XX  
 XX OS Synthetic.  
 XX XX  
 XX PN DE19816141-A1.  
 XX XX  
 XX PD 14-OCT-1999.  
 XX XX  
 XX PF 09-APR-1998; 98DE-01016141.  
 XX XX  
 XX PR 09-APR-1998; 98DE-01016141.  
 XX XX  
 XX PA (HMRI) HOECHST MARION ROUSSEL DEUT GMBH.  
 XX PA Kontermann R, Sedlacek H, Mueller R;  
 XX PI WPI; 1999-581511/50.  
 XX DR WPI; 1999-581511/50.  
 XX XX  
 XX PT New polypeptide binding agents containing variable heavy and light  
 XX PT constructs connected via peptide linker, used for treatment, prevention  
 XX PT or diagnosis of e.g. cancer.  
 XX XX  
 XX PS Claim 7; Page 17; 20pp; German.  
 XX XX  
 CC This sequence represents a novel single-chain molecule (I) that binds  
 CC multiple antigens and comprises two variable domains of heavy  
 CC immunoglobulin chains (VH), having specificities A and B and two variable  
 CC domains of light chains (VL), also with specificities A and B. The  
 CC domains are provided as two VH-VL constructs which are attached via a  
 CC peptide (P). Any VH and VL may be replaced by their functional fragments.  
 CC The products of the invention have anticancer, antiviral, antibacterial,  
 CC antimalarial and antiinflammatory activity. (I) are used to treat;  
 CC prevent or diagnose tumors (e.g. as tumor vaccines), autoimmune diseases  
 CC and inflammation (e.g. transplant rejection and arthritis), blood  
 CC disorders (e.g. of the coagulation and/or circulatory systems, such as  
 CC anemia, leucopenia, thrombocytopenia and hypertension), nervous system  
 CC disorders and/or infections (by viruses or bacteria, or malaria),  
 CC including, when (I) include a fusogenic peptide, use for gene transfer.  
 CC (I) are produced simply and in predominantly homogeneous form, in a wide  
 CC variety of hosts, either in secreted or membrane-bound forms. This  
 CC sequence represents a VH-VL domain linker peptide which is used to  
 CC illustrate the method of the invention

XX XX  
 XX SQ Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;  
 Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 DB 1 GGGGS 5

RESULT 9  
 AAY43496  
 ID AAY43496 strand; peptide; 5 AA.

AC AAY43496;

DT 26-JAN-2000 (first entry)

DE Linker for dual avb3 receptor/metastasis-associated receptor ligands.

Interferon-alpha-2b; IFN-alpha; avb3 antagonist; avb3 receptor ligand;  
 metastasis-associated receptor ligand; angiogenesis; cell proliferation;  
 anti-angiogenic protein; avb3-integrin; cancer; arthritis;  
 macular degeneration; diabetic retinopathy; hemangioma; psoriasis;  
 osteoporosis; thrombosis; angina; atherosclerosis; antiviral;  
 antibacterial; antifungal.

OS Homo sapiens.

PN WO951638-A1.

PD 14-OCT-1999.

PP 07-APR-1999; 99WO-US004295.

PR 08-APR-1998; 98US-0081074P.

PA (SEAR) SEARLE & CO G D.

PI Tjoeng FS, Fok KF;

DR WPI; 1999-620196/53.

PT New conjugates of integrin antagonist and ligand for metastasis-associated receptor, for treating angiogenesis-related diseases, e.g. cancer.

PS Claim 18; Page 86; 108pp; English.

The present sequence represents a linker used to join the avb3 antagonist and the metastasis-associated receptor ligand, in the pharmaceutical compounds of the invention. These compounds are dual avb3 receptor/metastasis-associated receptor ligands, and inhibit angiogenesis and thus proliferation of (cancer) cells. One component binds to the avb3 receptor and the other to a metastasis-associated receptor. The avb3 antagonists may also be conjugated to anti-angiogenic proteins, such as IFN-alpha and its derivatives. The compounds are used to treat angiogenesis-related disorders (mediated by the avb3-integrin), specifically cancer (of lung, breast, ovary, prostate, stomach, colon, kidney or bladder, also melanoma, hepatoma, sarcoma and lymphoma), arthritis and macular degeneration, and also diabetic retinopathy, hemangioma, psoriasis, osteoporosis, thrombosis, angina, atherosclerosis etc. The compounds may also be useful as antiviral, antibacterial and antifungal agents

CC Sequence 5 AA;

Query Match 100.0%; Score 28; DB 2; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 DB 1 GGGGS 5

RESULT 10

AAV83210  
 ID AAV83210 strand; peptide; 5 AA.

AC AAV83210;

DT 24-JUL-2000 (first entry)

DE Peptide linker used in construction of a\_vb\_3 integrin/IFN alpha.

Biconjugate; a\_vb\_3 integrin; interferon alpha; angiogenesis; cancer; tumor; osteoporosis; Paget's disease; Kaposi's sarcoma; periodontal disease; metastasis; neoplasia; retinopathy; arthritis; psoriasis; leukaemia; malignant melanoma; atherosclerosis; smooth muscle cell migration; inhibition; treatment; antagonist; angina; thrombosis; restenosis; antiviral; antifungal; antibacterial.

OS Synthetic.

PN WO200009143-A1.

PD 24-FEB-2000.

PP 07-APR-1999; 99WO-US004296.

PR 13-AUG-1998; 98US-0096442P.

PA (SEAR) SEARLE & CO G D.

PI Fok KF, Tjoeng FS;

DR WPI; 2000-205894/18.

PT New biconjugates comprising an avb3 antagonist and a metastatic-associated receptor ligand, useful for treating cancer and other angiogenic diseases, or as antiviral, antifungal or antibacterial agents.

PS Claim 19; Page 88; 123pp; English.

Biconjugates comprising one or more a\_vb\_3 antagonist moieties coupled to a peptide or polypeptide having anti-angiogenic properties can be used for treating a human patient with an angiogenesis-mediated disease, e.g. cancer, arthritis, or macular degeneration. The a\_vb\_3 integrin is normally associated with endothelial cells but can promote the formation of blood vessels (angiogenesis) in tumours. The a\_vb\_3 integrin is also known to play a role in tumour metastasis, neoplasia, osteoporosis, Paget's disease, retinopathy, arthritis, periodontal disease, psoriasis and smooth muscle cell migration. Interferon alpha is a family of proteins which possess complex antiviral, antineoplastic and immunomodulating activities. Interferon alpha is effective against a variety of cancers including hairy cell leukaemia, chronic myelogenous leukaemia, malignant melanoma and Kaposi's sarcoma. Multi-functional biconjugates comprising both a\_vb\_3 antagonists and interferon alpha 2b can exhibit greater biological activity when compared to a single factor or having improved half-life or decreased adverse side effects, or a combination of these properties. They can be used for inhibiting elevated levels of tumor antigens, inhibiting the proliferation of tumor cells and inhibiting tumor growth. The biconjugates can also be used for treating e.g. osteoporosis, humoral hypercalcemia of malignancy, Paget's disease, retinopathy including diabetic retinopathy, arthritis, including rheumatoid arthritis, periodontal disease, psoriasis, thrombosis, angina, atherosclerosis, smooth muscle cell migration and restenosis in a mammal. CC They are also useful as antiviral, antifungal, antibacterial agents. CC This sequence is a peptide linker used in the construction of the multi-functional biconjugates

CC Sequence 5 AA;

Query Match 100.0%; Score 28; DB 3; Length 5;

Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

## RESULT 11

AA06226 standard; peptide; 5 AA.

XX AAB06226;

DT 22-NOV-2000 (first entry)

DE Expression vector CANTAB5E inserted peptide.

XX Modified RNase; eosinophil derived neurotoxin protein; EDN; cancer;

KM Kaposi's sarcoma; neoplastic endothelial cell;

XX non-neoplastic endothelial cell; expression vector.

OS Synthetic.

PN W0200026233-A1.

PD 11-MAY-2000.

PF 01-NOV-1999; 99WC-US025737.

PR 02-NOV-1998; 98US-0106732P.

PA (USSH ) US DEPT HEALTH & HUMAN SERVICES.

PI Rybak SM, Newton DL;

DR WPI; 2000-365565/31.

XX N-terminally modified RNase A targeted to and are cytotoxic to cancerous

PT endothelial cells used to treat especially Kaposi's sarcoma.

XX Example 9; Page 34; 51pp; English.

XX The present sequence is a peptide which was inserted into expression

CC vector pCANTAB5E to enable more flexible folding of the human eosinophil

CC derived neurotoxin protein (EDN), which was expressed by the vector. This

CC protein can be directed to cancerous cells using additional N-terminal

CC peptides, where it exerts a cytotoxic effect. The protein can, therefore,

CC be used to treat cancer, particularly Kaposi's sarcoma, and to

CC selectively kill neoplastic and non-neoplastic endothelial cells

XX Sequence 5 AA;

## RESULT 12

AA54917 standard; peptide; 5 AA.

XX AAY54917;

DT 14-FEB-2000 (first entry)

DE Linker from IL-12 fusion protein.

XX Interleukin-12; IL-12; fusion protein; IL-12 p35 subunit; B7 protein;

KM IL-12 p40 subunit; gene therapy; tumour; leukaemia; linker.

OS Synthetic.

XX US5994104-A.

XX 30-NOV-1999.

PF 08-NOV-1996; 96US-00751767.

PR 08-NOV-1996; 96US-00751767.

PA (UNLO ) ROYAL FREE HOSPITAL SCHOOL MED.

PI Anderson RJ, Prentice HG, Macdonald ID;

DR WPI; 2000-038261/03.

XX Nucleic acid constructs encoding interleukin-12 fusion proteins useful

PT for treating leukemia and other cancers.

PS Claim 3; Col 93; 73pp; English.

XX This sequence represents a linker that can be used in an interleukin-12

CC fusion protein. The invention relates to an isolated nucleic acid

CC construct (1) comprising a region encoding an interleukin-12 (IL-12)

CC fusion protein (comprising an IL-12 p35 subunit, an IL-12 p40 subunit and

CC a linker peptide (joining the subunits)) and a region encoding a B7

CC protein. (1) may be used to produce IL-12 fusion proteins according to

CC standard recombinant DNA methodologies. The fusion proteins may be

CC produced either in vitro in a fermentation culture or in vivo as part of

CC a gene therapy protocol (in this case (1) is used to transform a patient's

CC cells, which then secrete the functional polypeptide to supplement the

CC patient's own production of IL-12 or to rectify mutations which lead to

CC the expression of inactive polypeptides). The fusion proteins produced in

CC this way may be used to treat any disease which responds to IL-12 such as

CC tumours (both solid and dispersed of the kidney, breast, colon, ovarian

CC and cervical tumours and melanomas) and in particular, tumours of the

CC blood such as leukaemia. Alternatively, the polypeptides may be used as

CC agonists and antagonists of its activity. The antibodies and antagonists

CC may be used to inhibit the activity of IL-12. (1) may also be used

CC diagnostically as a probe which hybridizes to sequences encoding IL-12

CC and the antibodies may be used to detect the presence of IL-12

CC polypeptides in samples. They may be used diagnostically to quantitate

CC the expression of the polypeptide by patients and hence which subjects

CC may be in need of restorative therapy

XX Sequence 5 AA;

XX Query Match 100.0%; Score 28; DB 3; Length 5;

XX Best Local Similarity 100.0%; Pred. No. 1.7e+06;

XX Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGS 5  
 |||||  
 DB 1 GGGGS 5

AA43750 standard; peptide; 5 AA.

XX AAY43750;

DT 11-FEB-2000 (first entry)

DE Linker used to construct a bispecific single-chain antibody.

XX bisCD19xCD3 antibody; bispecific single-chain fragment; CD19 antigen;

KM CD3 antigen; CD19-positive target cell; T-cell stimulation;

KM cytotoxic T-lymphocyte; B-cell malignancy; myasthenia gravis;

KM B-cell mediated autoimmune disease; Morbus Basedow;

KM Hashimoto thyroiditis; Goodpasture syndrome; B-cell depletion;

KM non-Hodgkin lymphoma; gene therapy; cancer; viral disease.

XX Synthetic.  
OS  
XX  
XX WO9554440-A1.  
XX  
XX 28-OCT-1999.  
XX  
XX 21-APR-1999; 99WO-EP002693.  
XX  
XX 21-APR-1998; 98EP-00107269.  
XX  
XX (DOER/) DOERKEN B.  
XX (RIET/) RIETHMUELLER G.  
XX  
XX Kufer P, Luterbuese R, Bargou R, Loeffler A;  
XX WPI; 2000-013241/01.  
XX  
XX Novel multifunctional polypeptide for treating B-cell malignancies  
XX especially non-Hodgkin lymphoma.  
XX  
XX Claim 10; Page 49; 91pp; English.  
XX  
XX The present sequence represents a linker used in the construction of  
XX bi-specific single-chain polypeptides of the invention. These polypeptides  
XX comprise domains providing binding-site of immunoglobulin chains or  
XX antibodies specifically recognizing CD19 and CD3 antigen. The polypeptide  
XX destroys CD19-positive target cells without any need of T-cell pre and/or  
XX co-stimulation, by recruiting cytotoxic T-lymphocytes and so specific  
XX lysis by T-cells rather than a direct effect by an antibody is achieved.  
XX The bi-specific single-chain polypeptides, or nucleotides encoding them,  
XX are used for the treatment of B-cell malignancies, B-cell mediated  
XX autoimmune diseases like myasthenia gravis, Morbus Basedow, Hashimoto  
XX thyroiditis or Goodpasture syndrome or for the depletion of B-cells and  
XX more particularly non-Hodgkin lymphoma in mammals preferably human. They  
XX can also delay the pathological conditions caused by these diseases, and  
XX can be used for detecting these diseases. The polynucleotide is used for  
XX gene therapy. The polypeptides are also used for identifying compounds  
XX modulating B-cell/T-cell mediated immune response with can in turn be  
XX used for treating cancer, its related diseases and also for inhibiting  
XX viral diseases by preventing viral infection  
XX  
XX Sequence 5 AA;  
SQ

Query Match 100.0%; Score 28; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 14  
AAB14535  
ID AAB14535 standard; peptide; 5 AA.  
XX  
XX AAB14535;  
XX  
XX 24-NOV-2000 (first entry)  
XX  
XX Peptide linker for joining HIV-1 gp41 N- and C-terminal helices.  
XX  
XX HIV-1; gp41; N-helical domain; heptad repeat region; C-helical domain;  
XX gp41 transmembrane-proximal amphipathic alpha-helical segment;  
XX core 6-helix bundle; viral entry inhibition; immunogenic; antibody;  
XX humoral response; broad spectrum vaccine; anti-HIV.  
XX envelope glycoprotein; prophylaxis; therapy; peptide linker.  
XX  
XX Synthetic.  
XX  
XX WO2000040616-A1.  
XX  
XX

PD 13-JUL-2000.  
XX  
XX 10-JAN-2000; 2000WO-US000456.  
XX  
XX 08-JAN-1999; 99US-0115404P.  
XX 07-JAN-2000; 2000US-00480336.  
XX  
XX (WILD/) WILD C.T.  
XX (WEIS/) WEISS C.D.  
XX  
XX Wild CT, Weiss CD;  
XX  
XX WPI; 2000-465959/40.  
XX  
XX Raising neutralizing antibody response to human immunodeficiency virus,  
XX comprises administering a polypeptide capable of forming a stable coiled-  
XX coil solution structure.  
XX  
XX Disclosure; Page 15; 97pp; English.  
XX  
XX The invention relates to raising a neutralizing antibody response to a  
XX broad spectrum of HIV (human immunodeficiency virus) strains and  
XX isolates, comprising the administration of a peptide which corresponds to  
XX or mimics highly conserved portions of the gp41 envelope glycoprotein  
XX which are important in mediating the process of viral entry into host  
XX cells. Such peptides can correspond to or mimic the coiled coil solution  
XX structure of the N-helical domain (the heptad repeat region), or can  
XX correspond or mimic the C-helical domain (the transmembrane-proximal  
XX amphipathic alpha-helical segment), or the gp41 core 6-helix bundle,  
XX which is formed by the interaction of the N- and C-helical domains of  
XX three gp41 proteins. The peptides can be administered either singly or as  
XX a combination (particularly a combination of N-helical and C-helical  
XX peptides), and can be multimerised. For example, N- and C-helical domain  
XX peptides can be alternately linked together to form a peptide(s)  
XX mimics the core 6-helix bundle. Administration of the peptide(s)  
XX generates a humoral response, with the production of antibodies against  
XX gp41 structures involved in viral entry. As these portions of gp41 are  
XX well conserved, such antibodies may be effective against a broad range of  
XX HIV strains and isolates. The peptide compositions may be administered as  
XX a prophylactic or therapeutic vaccine to generate antibodies which reduce  
XX or inhibit the ability of HIV to infect uninfected cells. A composition  
XX comprising polyclonal or monoclonal antibodies can be administered to  
XX reduce HIV infection of uninfected cells. Antibodies raised against entry  
XX -relevant gp41 structures may also be used therapeutically and as tools  
XX to further elucidate the mechanism of HIV cell entry. The present  
XX sequence represents a peptide linker which may be used to join peptides  
XX of the invention together to form multimers  
XX  
XX Sequence 5 AA;  
SQ

Query Match 100.0%; Score 28; DB 3; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.7e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 15  
AAB00156  
ID AAB00156 standard; peptide; 5 AA.  
XX  
XX AAB00156;  
XX  
XX 08-FEB-2001 (first entry)  
XX  
XX Linker used in scd4-SCFv(17b) fusion protein.  
XX  
XX Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
XX acquired immune deficiency syndrome; neutralisation; infection;  
XX gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
XX binding domain; single chain antibody; chimera; chimeric protein.  
XX  
XX

```

XX      Synthetic.
OS
XX      WO200055207-A1.
PN
XX      21-SEP-2000.
PD
XX      16-MAR-2000; 2000WO-US006946.
PF
XX      16-MAR-1999; 99US-0124681P.
PR
XX      (USSH ) US NAT INST OF HEALTH.
PA
XX      Berger EA, Del Castillo CM;
PI
XX      WPI; 2000-638183/61.
DR
XX      Novel neutralizing bispecific fusion proteins effective in viral such as
PT      HIV neutralization, comprises two different binding domains, inducing-
PT      binding domain and induced-binding domain functionally linked by linker.
XX
XX      Claim 30; Page 45; 55pp; English.
PS
XX
XX      bCD4-SCFv(17b) is a neutralizing bispecific fusion protein capable of
CC      binding to two sites of its target protein. The protein comprises a first
CC      binding domain capable of binding to an inducing site on the target
CC      protein, a second binding domain capable of forming neutralising complex
CC      with an induced epitope of the target protein and a linker connecting the
CC      binding domains. bCD4-SCFv(17b) comprises a soluble CD4 fragment
CC      (containing domains D1 and D2) fused to a single chain Fv portion of
CC      antibody 17b via a linker. bCD4-SCFv(17b), its variant, analogue or
CC      mimetic is used for inactivating gp120 protein of HIV, and for
CC      neutralising HIV. It is also used for blocking and preventing the binding
CC      of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte
CC      CD4 and for inhibiting HIV replication. The chimeric proteins is
CC      therefore useful for treating HIV infection and also AIDS. It is are
CC      particularly useful in the prevention of infection during or immediately
CC      after HIV exposure (e.g., mother/infant transmission, post-exposure
CC      prophylaxis, and as a topical inhibitor) and for providing long term
CC      resistance to HIV infections and AIDS. Gene therapy is used to secrete
CC      the bispecific protein at mucosal surfaces, such as the vaginal, rectal
CC      or oral mucosa. The fusion proteins is highly potent, broadly cross-
CC      reactive with neutralising antibody with high in vivo activity and no FC-
CC      mediated undesirable targeting properties. When the fusion protein is
CC      substantially derived from human proteins, it has minimal immunogenicity
CC      and toxicity in humans which is of great value in prevention of infection
CC      during or immediately after HIV exposure
XX
XX      Sequence 5 AA;
SQ

```

```

Query Match          100.0%; Score 28; DB 3; Length 5;
Best Local Similarity 100.0%; Pred. NO. 1.7e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      1 GGGGS 5
        |||||
Db      1 GGGGS 5

```

Search completed: November 21, 2004, 13:37:11  
 Job time : 5.58212 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:56 ; Search time 0.64781 Seconds  
(without alignments)  
511.863 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 478139 seqs, 66318000 residues

Total number of hits satisfying chosen parameters: 478139

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database : Issued Patents AA:  
1: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/1/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/prodata/1/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/prodata/1/1aa/5C\_COMB.pep:\*  
6: /cgn2\_6/prodata/1/1aa/5D\_COMB.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	28	100.0	5 1 US-07-959-946-6	Sequence 6, Appl
2	28	100.0	5 1 US-08-176-500-140	Sequence 140, App
3	28	100.0	5 1 US-08-471-052A-140	Sequence 140, App
4	28	100.0	5 1 US-08-225-224-54	Sequence 54, Appl
5	28	100.0	5 1 US-08-236-918A-18	Sequence 18, Appl
6	28	100.0	5 1 US-08-463-163-1	Sequence 1, Appl
7	28	100.0	5 1 US-08-566-800A-58	Sequence 58, Appl
8	28	100.0	5 1 US-08-244-469-5	Sequence 5, Appl
9	28	100.0	5 1 US-08-189-331-140	Sequence 140, App
10	28	100.0	5 1 US-08-333-577-6	Sequence 6, Appl
11	28	100.0	5 1 US-08-575-361A-32	Sequence 32, Appl
12	28	100.0	5 2 US-08-564-955-64	Sequence 64, Appl
13	28	100.0	5 2 US-08-528-523-13	Sequence 13, Appl
14	28	100.0	5 2 US-08-537-874-62	Sequence 62, Appl
15	28	100.0	5 2 US-08-448-418-86	Sequence 86, Appl
16	28	100.0	5 2 US-08-471-939-140	Sequence 140, App
17	28	100.0	5 2 US-08-471-800-140	Sequence 140, App
18	28	100.0	5 2 US-08-932-589-58	Sequence 58, Appl
19	28	100.0	5 2 US-08-647-449-26	Sequence 26, Appl
20	28	100.0	5 2 US-08-821-840-1	Sequence 1, Appl
21	28	100.0	5 2 US-08-902-623-75	Sequence 75, Appl
22	28	100.0	5 2 US-08-471-068-140	Sequence 140, App
23	28	100.0	5 2 US-08-308-494A-1	Sequence 1, Appl
24	28	100.0	5 2 US-08-665-202-25	Sequence 25, Appl
25	28	100.0	5 2 US-08-809-668-12	Sequence 12, Appl
26	28	100.0	5 2 US-08-792-553-15	Sequence 15, Appl
27	28	100.0	5 2 US-08-751-767A-24	Sequence 24, Appl

28	28	100.0	5 2 US-08-818-253-40	Sequence 40, Appl
29	28	100.0	5 3 US-08-722-258-54	Sequence 54, Appl
30	28	100.0	5 3 US-09-397-951-12	Sequence 12, Appl
31	28	100.0	5 3 US-08-776-271-3	Sequence 3, Appl
32	28	100.0	5 3 US-08-621-859-64	Sequence 64, Appl
33	28	100.0	5 3 US-09-100-856A-62	Sequence 62, Appl
34	28	100.0	5 3 US-09-046-992-5	Sequence 5, Appl
35	28	100.0	5 3 US-09-215-035-3	Sequence 3, Appl
36	28	100.0	5 3 US-09-075-511-64	Sequence 64, Appl
37	28	100.0	5 3 US-09-099-015-64	Sequence 64, Appl
38	28	100.0	5 3 US-08-818-252-40	Sequence 40, Appl
39	28	100.0	5 3 US-09-091-814-1	Sequence 1, Appl
40	28	100.0	5 3 US-09-281-792B-26	Sequence 26, Appl
41	28	100.0	5 3 US-09-232-863-62	Sequence 62, Appl
42	28	100.0	5 3 US-09-362-805-1	Sequence 1, Appl
43	28	100.0	5 3 US-09-362-805-2	Sequence 2, Appl
44	28	100.0	5 3 US-09-367-953B-26	Sequence 26, Appl
45	28	100.0	5 3 US-09-133-508A-62	Sequence 62, Appl

## ALIGNMENTS

RESULT 1  
US-07-959-946-6  
; Sequence 6, Application US/07959946  
; Patent No. 5408038  
; GENERAL INFORMATION:  
; APPLICANT: Smith, Richard K.  
; APPLICANT: Koduri, Raju  
; APPLICANT: Young, Stephen G.  
; APPLICANT: Wilczum, Joseph L.  
; APPLICANT: Cartles, Linda K.  
; TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
; NUMBER OF SEQUENCES: 20  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Dressler, Goldsmith, Shore, Sucker &  
; ADDRESSEE: Milnamow, Ltd.  
; STREET: 180 No. 5408038th Stetson, Suite 4700  
; CITY: Chicago  
; STATE: Illinois  
; COUNTRY: USA  
; ZIP: 60601  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/07/959, 946  
; CLASSIFICATION: 435  
; FILING DATE: 19921008  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/901,706  
; FILING DATE: 18-JUN-1992  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Gamson, Edward P.  
; REGISTRATION NUMBER: 29,381  
; REFERENCE/DOCKET NUMBER:  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (312)616-5400  
; TELEFAX: (312)616-5460  
; INFORMATION FOR SEQ ID NO: 6:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-07-959-946-6  
Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 2

US-08-176-500-140  
; Sequence 140, Application US/08176500  
; Patent No. 5498538  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowlkes, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 141  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/176,500  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/08/013,416  
; FILING DATE:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-143  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 140:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-176-500-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 3

US-08-471-052A-140  
; Sequence 140, Application US/08471052A  
; Patent No. 5625033  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowlkes, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 166  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas

CITY: New York  
STATE: New York  
COUNTRY: U.S.A.  
ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/471,052A  
; FILING DATE: 06-JUNE-1995  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mirock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-179  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8864/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 140:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
; US-08-471-052A-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

## RESULT 4

US-08-225-224-54  
; Sequence 54, Application US/08225224  
; Patent No. 5635599  
; GENERAL INFORMATION:  
; APPLICANT: PASTAN, Ira  
; APPLICANT: KREITMAN, Robert J.  
; TITLE OF INVENTION: CIRCULARLY PERMUTATED LIGANDS AND  
; TITLE OF INVENTION: CIRCULARLY PERMUTED FUSION PROTEINS  
; NUMBER OF SEQUENCES: 57  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Townsend and Townsend Hourie and Crew  
; STREET: Stewart Street Tower, One Market Plaza  
; CITY: San Francisco  
; STATE: California  
; COUNTRY: US  
; ZIP: 94105-1493  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/225,224  
; FILING DATE: 8-APR-1994  
; CLASSIFICATION: 530  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Weber, Ellen L.  
; REGISTRATION NUMBER: 32,762  
; REFERENCE/DOCKET NUMBER: 15280-193  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 543-9600  
; TELEFAX: (415) 543-5043  
; INFORMATION FOR SEQ ID NO: 54:

SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: unknown  
TOPOLOGY: unknown  
MOLECULE TYPE: peptide  
US-08-225-224-54

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 5  
US-08-236-918A-18  
Sequence 18, Application US/08236918A  
Patent No. 5674704  
GENERAL INFORMATION:  
APPLICANT: Alderson, Mark R.  
APPLICANT: Goodwin, Raymond G.  
APPLICANT: Smith, Craig A.  
TITLE OF INVENTION: Cytokine Designated 4-1BB Ligand  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Kathryn A. Anderson, Immunex Corporation  
STREET: 51 University Street  
CITY: Seattle  
STATE: Washington  
COUNTRY: US  
ZIP: 98101  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: Apple Power Macintosh  
OPERATING SYSTEM: Apple 7.5.3  
SOFTWARE: Microsoft Word, Version #6.0.1  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/236,918A  
FILING DATE: 06-May-1994  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/060,843  
FILING DATE: 07-May-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Anderson, Kathryn A.  
REGISTRATION NUMBER: 32,172  
REFERENCE/DOCKET NUMBER: 2801-B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (206) 233-0644  
TELEFAX: (206) 233-0644  
INFORMATION FOR SEQ ID NO: 18:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-236-918A-18

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 6  
US-08-463-163-1

Sequence 1, Application US/08463163  
Patent No. 5696237  
GENERAL INFORMATION:  
APPLICANT: FitzGerald, David J.  
APPLICANT: Chaudhary, Vijay K.  
APPLICANT: Pastan, Ira H.  
APPLICANT: Waldmann, Thomas A.  
APPLICANT: Queen, Cary L.  
TITLE OF INVENTION: Recombinant Antibody-Toxin Fusion Protein  
NUMBER OF SEQUENCES: 12  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew  
STREET: One Market Plaza, Stewart Street Tower  
CITY: San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94105-1492  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,163  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/227,227  
FILING DATE: 22-JAN-1981  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 06/911,227  
FILING DATE: 24-SEP-1986  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/341,361  
FILING DATE: 21-APR-1989  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/865,722  
FILING DATE: 08-APR-1992  
ATTORNEY/AGENT INFORMATION:  
NAME: Weber, Ellen L.  
REGISTRATION NUMBER: 32,762  
REFERENCE/DOCKET NUMBER: 015280-12211  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 543-9600  
TELEFAX: (415) 543-5043  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-08-463-163-1

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 7  
US-08-566-800A-58  
Sequence 58, Application US/08566800A  
Patent No. 5736364  
GENERAL INFORMATION:  
APPLICANT: Kelley, Robert F.  
APPLICANT: Lazarus, Robert A.  
APPLICANT: Lee, Geoffrey F.  
TITLE OF INVENTION: No. 5736364e1 Factor VIIa Inhibitors  
NUMBER OF SEQUENCES: 58

CORRESPONDENCE ADDRESS:  
ADDRESSEE: Genentech, Inc.  
STREET: 460 Point San Bruno Blvd  
CITY: South San Francisco  
STATE: California  
COUNTRY: USA  
ZIP: 94080  
COMPUTER READABLE FORM:  
MEDIUM TYPE: 3.5 inch, 1.44 Mb floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: WinPacIn (Genentech)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/566,800A  
FILING DATE: 04-Dec-1995  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kurbinec, Jeffrey S.  
REGISTRATION NUMBER: 36,575  
REFERENCE/DOCKET NUMBER: P0958B  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415/225-8228  
TELEFAX: 415/952-9881  
TELEX: 910/371-7168  
INFORMATION FOR SEQ ID NO: 58:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: Amino Acid  
TOPOLOGY: Linear  
US-08-566-800A-58

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 8  
US-08-244-469-5  
; Sequence 5, Application US/08244469  
; Patent No. 5736387  
; GENERAL INFORMATION:  
; APPLICANT: Paul, Ralph W.  
; APPLICANT: Overell, Robert  
; TITLE OF INVENTION: ENVELOPE FUSION VECTORS FOR USE IN GENE  
; TITLE OF INVENTION: DELIVERY  
; NUMBER OF SEQUENCES: 7  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: MORRISON & FOERSTER  
; STREET: 755 PAGE MILL ROAD  
; CITY: PALO ALTO  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94304-1018  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/244,469  
; FILING DATE: 01-JUN-1994  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Dylan, Tyler M.  
; REGISTRATION NUMBER: 37,612  
; REFERENCE/DOCKET NUMBER: 22627-20007.20  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (415) 813-5600  
; TELEFAX: (415) 494-0792

TELEX: 706141 MRSNPOERS SFO  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-244-469-5

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 9  
US-08-189-331-140  
; Sequence 140, Application US/08189331  
; Patent No. 5747334  
; GENERAL INFORMATION:  
; APPLICANT: Kay, B. K.  
; APPLICANT: Fowlkes, D. M.  
; TITLE OF INVENTION: Totally Synthetic Affinity Reagents  
; NUMBER OF SEQUENCES: 186  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: New York  
; COUNTRY: U.S.A.  
; ZIP: 10036-2711  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/189,331  
; FILING DATE: Concurrently herewith  
; CLASSIFICATION: 435  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mierock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 1101-155  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 212 790-9090  
; TELEFAX: 212 869-8664/9741  
; TELEX: 66141 PENNIE  
; INFORMATION FOR SEQ ID NO: 140:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 5 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: unknown  
; MOLECULE TYPE: peptide  
US-08-189-331-140

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 10  
US-08-333-577-6  
; Sequence 6, Application US/08333577  
; Patent No. 5786206

GENERAL INFORMATION:  
APPLICANT: Smith, Richard K.  
APPLICANT: Koduri, Raju  
APPLICANT: Young, Stephen G.  
APPLICANT: Wiltum, Joseph L.  
APPLICANT: Curtiss, Linda K.  
TITLE OF INVENTION: Lipoprotein Assays Using Antibodies to a  
TITLE OF INVENTION: Pan Native Epitope and Recombinant Antigens  
NUMBER OF SEQUENCES: 20  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Dressler, Goldsmith, Shore, Sutker &  
ADDRESSEE: Milnamow, Ltd.  
STREET: 180 No. 5786206th Stetson, Suite 4700  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60601  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/333,577  
FILING DATE:  
CLASSIFICATION: 530  
ATTORNEY/AGENT INFORMATION:  
NAME: Gamson, Edward P.  
REGISTRATION NUMBER: 29,381  
REFERENCE/DOCKET NUMBER: SCRP 234.0  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (312)616-5400  
TELEFAX: (312)616-5460  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-333-577-6

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
DB 1 GGGGS 5

RESULT 11  
US-08-575-361A-32  
Sequence 32, Application US/08575361A  
Patent No. 5792640  
GENERAL INFORMATION:  
APPLICANT: Chandrasegaran, Srinivasan  
TITLE OF INVENTION: A GENERAL METHOD TO CLONE HYBRID  
TITLE OF INVENTION: RESTRICTION ENDONUCLEASES USING 11g GENE  
NUMBER OF SEQUENCES: 35  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Cushman Darby & Cushman L.L.P.  
STREET: 1100 New York Avenue, NW, Ninth Floor, East  
STREET: Tower  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20005-3918  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/575,361A  
FILING DATE: 20-DEC-1995  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Kokulig, Paul N.  
REGISTRATION NUMBER: 16,773  
REFERENCE/DOCKET NUMBER: PNK/4130/213779/DJP  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-861-3000  
TELEFAX: 202-822-0944  
INFORMATION FOR SEQ ID NO: 32:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
US-08-575-361A-32

Query Match 100.0%; Score 28; DB 1; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
DB 1 GGGGS 5

RESULT 12  
US-08-564-955-64  
Sequence 64, Application US/08564955  
Patent No. 5811238  
GENERAL INFORMATION:  
APPLICANT: STEMMER, WILLEM P.C.  
APPLICANT: CRAMERI, ANDREAS M.  
TITLE OF INVENTION: METHODS FOR GENERATING POLYNUCLEOTIDES  
TITLE OF INVENTION: HAVING DESIRED CHARACTERISTICS BY ITERATIVE SELECTION AND  
TITLE OF INVENTION: RECOMINATION  
NUMBER OF SEQUENCES: 67  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: TOWNSEND AND TOWNSEND AND CREW  
STREET: TWO EMBARCADERO CENTER, 8TH FLOOR  
CITY: SAN FRANCISCO  
STATE: CALIFORNIA  
COUNTRY: U.S.A.  
ZIP: 94111-3834  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/564,955  
FILING DATE: 30-NOV-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/198,431  
FILING DATE: 17-FEB-1994  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 08/537,874  
FILING DATE: 30-OCT-1995  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/US95/02126  
FILING DATE: 17-FEB-1995  
ATTORNEY/AGENT INFORMATION:  
NAME: DUNN, TRACY J.  
REGISTRATION NUMBER: 34,587  
REFERENCE/DOCKET NUMBER: 165261-014611US  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (415) 326-2400  
TELEFAX: (415) 576-0300  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids

TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
US-08-564-955-64

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 13  
US-08-528-523-13  
Sequence 13, Application US/08528523  
Patent No. 5824782  
GENERAL INFORMATION:  
APPLICANT: Hoeizer, Wolfgang  
APPLICANT: von Hoegen, Ilka  
APPLICANT: Strittmatter, Wolfgang  
APPLICANT: Matzku, Siegfried  
TITLE OF INVENTION: Immunconjugates II  
NUMBER OF SEQUENCES: 13  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Millen, White, Zelano & Branigan, P.C.  
STREET: 2200 Clarendon Boulevard, Suite 1400  
CITY: Arlington  
STATE: Virginia  
COUNTRY: U.S.A.  
ZIP: 22201  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/528,523  
FILING DATE: 06-NOV-1992  
CLASSIFICATION: 536  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 94114572.4  
FILING DATE: 16-SEP-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Hamlet-King, Diana  
REGISTRATION NUMBER: 33,302  
REFERENCE/DOCKET NUMBER: Merck 1717  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 703-243-6333  
TELEFAX: 703-243-6410  
TELEX: 64191  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
HYPOTHEICAL: NO  
FRAGMENT TYPE: internal  
US-08-528-523-13

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 14

US-08-537-874-62  
Sequence 62, Application US/08537874  
Patent No. 5830721  
GENERAL INFORMATION:  
APPLICANT: Stemmer, Willem P.C.  
APPLICANT: Cramer, Andreas  
TITLE OF INVENTION: DNA Mutagenesis by Random Fragmentation  
TITLE OF INVENTION: and Reassembly  
NUMBER OF SEQUENCES: 62  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Townsend and Townsend and Crew LLP  
STREET: Two Embarcadero Center, 8th Floor  
CITY: San Francisco  
STATE: CA  
COUNTRY: USA  
ZIP: 94111

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/537,874  
FILING DATE:  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP PCT/US95/02126  
FILING DATE: 17-FEB-1995  
APPLICATION NUMBER: US 08/198,431  
FILING DATE: 17-FEB-1994  
ATTORNEY/AGENT INFORMATION:  
NAME: Liebeschuetz, Joe  
REGISTRATION NUMBER: 37,505  
REFERENCE/DOCKET NUMBER: 018097-014610  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 415-576-0200  
TELEFAX: 415-576-0300

INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
STRANDEDNESS: not relevant  
TOPOLOGY: not relevant  
US-08-537-874-62

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 15  
US-08-448-418-86  
Sequence 86, Application US/08448418  
Patent No. 5837242  
GENERAL INFORMATION:  
APPLICANT: Holliger, Kaapar-Philipp  
APPLICANT: Griffiths, Andrew D  
APPLICANT: Hoogenboom, Hendricus RJM  
APPLICANT: Malmqvist, Magnus  
APPLICANT: Marks, James D  
APPLICANT: McGuinness, Brian T  
APPLICANT: Pope, Anthony R  
APPLICANT: Prospero, Terence D  
APPLICANT: Winer, Gregory P  
TITLE OF INVENTION: Multivalent and Multispecific Binding  
TITLE OF INVENTION: Proteins, Their Manufacture and Use  
NUMBER OF SEQUENCES: 106  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Marshall O'Toole Gerstein Murray and Borun

STREET: 6300 Sears Tower 233 South Wacker Drive  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25 (EPO)  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/448,418  
FILING DATE: 14-MAY-1996  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA: C12N 15/62, 15/70, C07K 1/00  
APPLICATION NUMBER: PCT/GB93/02492  
FILING DATE: 03-DEC-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9225453.1  
FILING DATE: 04-DEC-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 9300816.7  
FILING DATE: 16-JAN-1993  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: EP 93303614.7  
FILING DATE: 10-MAY-1993  
PRIOR APPLICATION DATA: GB 9319969.3  
APPLICATION NUMBER: GB 9319969.3  
FILING DATE: 22-SEP-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: David W. Clough  
REGISTRATION NUMBER: 36,107  
REFERENCE/DOCKET NUMBER: 28111/32651  
INFORMATION FOR SEQ ID NO: 86:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: peptide linker  
US-08-448-418-86

Query Match 100.0%; Score 28; DB 2; Length 5;  
Best Local Similarity 100.0%; Pred. No. 3.8e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGS 5  
|||||  
Db 1 GGGS 5

Search completed: November 21, 2004, 13:44:43  
Job time: 1.64781 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 ; Search time 2.18978 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

```
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US10E_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	5	US-09-832-297A-1	Sequence 1, Appli
2	28	100.0	5	US-09-832-297A-2	Sequence 2, Appli
3	28	100.0	5	US-09-287-849-44	Sequence 44, Appl
4	28	100.0	5	US-09-147-142-31	Sequence 31, Appl
5	28	100.0	5	US-09-214-645-1	Sequence 1, Appli
6	28	100.0	5	US-09-858-616-2	Sequence 2, Appli
7	28	100.0	5	US-09-779-233-45	Sequence 45, Appl
8	28	100.0	5	US-09-989-789-3	Sequence 3, Appli
9	28	100.0	5	US-09-976-787-21	Sequence 21, Appl
10	28	100.0	5	US-09-192-854-180	Sequence 180, App
11	28	100.0	5	US-09-761-962-36	Sequence 36, Appl
12	28	100.0	5	US-09-333-527-5	Sequence 5, Appli
13	28	100.0	5	US-09-925-796-8	Sequence 8, Appli

14	28	100.0	5	US-09-815-837-116	Sequence 116, App
15	28	100.0	5	US-09-033-525-5	Sequence 5, Appli
16	28	100.0	5	US-09-779-451-7	Sequence 7, Appli
17	28	100.0	5	US-09-941-450-8	Sequence 8, Appli
18	28	100.0	5	US-09-818-247-25	Sequence 25, Appl
19	28	100.0	5	US-09-883-777-10	Sequence 10, Appl
20	28	100.0	5	US-09-867-262-3	Sequence 3, Appli
21	28	100.0	5	US-09-780-933-22	Sequence 22, Appl
22	28	100.0	5	US-09-480-238-10	Sequence 10, Appl
23	28	100.0	5	US-09-731-558-6	Sequence 6, Appli
24	28	100.0	5	US-09-828-708-123	Sequence 123, App
25	28	100.0	5	US-09-885-551A-3	Sequence 3, Appli
26	28	100.0	5	US-09-756-283A-14	Sequence 14, Appl
27	28	100.0	5	US-09-144-886-4	Sequence 4, Appli
28	28	100.0	5	US-09-999-745-56	Sequence 56, Appli
29	28	100.0	5	US-09-942-087A-8	Sequence 8, Appli
30	28	100.0	5	US-09-942-090-8	Sequence 8, Appli
31	28	100.0	5	US-09-554-000-40	Sequence 40, Appl
32	28	100.0	5	US-09-792-793A-1	Sequence 1, Appli
33	28	100.0	5	US-09-792-793A-2	Sequence 2, Appli
34	28	100.0	5	US-09-846-033B-212	Sequence 212, App
35	28	100.0	5	US-09-990-186-3	Sequence 3, Appli
36	28	100.0	5	US-09-897-844-8	Sequence 8, Appli
37	28	100.0	5	US-09-989-994-3	Sequence 3, Appli
38	28	100.0	5	US-09-911-261A-23	Sequence 23, Appli
39	28	100.0	5	US-09-942-024-84	Sequence 84, Appl
40	28	100.0	5	US-09-942-098-84	Sequence 84, Appl
41	28	100.0	5	US-09-969-748C-38	Sequence 38, Appl
42	28	100.0	5	US-09-992-124A-61	Sequence 61, Appl
43	28	100.0	5	US-09-949-039-37	Sequence 37, Appl
44	28	100.0	5	US-10-087-426-3	Sequence 3, Appli
45	28	100.0	5	US-10-057-505-15	Sequence 15, Appl

## ALIGNMENTS

```
RESULT 1
US-09-832-297A-1
; Sequence 1, Application US/09832297A
; Publication No. US20010055566A1
; GENERAL INFORMATION:
; APPLICANT: LUTKEN, George A.
; TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
; FILE REFERENCE: FLTOR1120-2
; CURRENT APPLICATION NUMBER: US/09/832,297A
; PRIOR FILING DATE: 2001-04-09
; PRIOR APPLICATION NUMBER: US 09/362,805
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: US 09/173,190
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 12
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 5
; TYPE: PRT
; ORGANISM: Artificial sequence
; FEATURE:
; OTHER INFORMATION: Peptide linker moiety
US-09-832-297A-1

Query Match      100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 2  
US-09-832-297A-2

```
/ Sequence 2, Application US/09832297A
/ Publication No. US20010055566A1
/ GENERAL INFORMATION:
/ APPLICANT: Fluoroprobe, Inc.
/ APPLICANT: LUTKEN, George A.
/ TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
/ FILE REFERENCE: FLUOR1120-2
/ CURRENT APPLICATION NUMBER: US/09/832,297A
/ PRIOR FILING DATE: 2001-04-09
/ PRIOR APPLICATION NUMBER: US 09/362,805
/ PRIOR FILING DATE: 1999-07-28
/ PRIOR APPLICATION NUMBER: US 09/173,190
/ PRIOR FILING DATE: 1998-10-15
/ NUMBER OF SEQ ID NOS: 12
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 2
/ LENGTH: 5
/ TYPE: PRT
/ ORGANISM: Artificial sequence
/ FEATURE:
/ OTHER INFORMATION: Peptide linker moiety. The sequence can be repeated n number of
US-09-832-297A-2
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 3
US-09-287-849-44
/ Sequence 44, Application US/09287849
/ Patent No. US2002009459A1
/ GENERAL INFORMATION:
/ APPLICANT: Reed, Steven G.
/ APPLICANT: Skeiky, Yasir A.W.
/ APPLICANT: Dillon, Davin C.
/ APPLICANT: Alderson, Mark
/ APPLICANT: Campos-Neto, Antonio
/ APPLICANT: Corixa Corporation
/ TITLE OF INVENTION: Fusion Proteins of Mycobacterium tuberculosis Antigens
/ TITLE OF INVENTION: and Their Uses
/ FILE REFERENCE: 014058-009020US
/ CURRENT APPLICATION NUMBER: US/09/287,849
/ PRIOR FILING DATE: 1999-04-07
/ PRIOR APPLICATION NUMBER: US 08/818,112
/ PRIOR FILING DATE: 1997-03-13
/ PRIOR APPLICATION NUMBER: US 08/942,578
/ PRIOR FILING DATE: 1997-10-01
/ PRIOR APPLICATION NUMBER: US 09/025,197
/ PRIOR FILING DATE: 1998-02-18
/ PRIOR APPLICATION NUMBER: US 09/056,556
/ PRIOR FILING DATE: 1998-04-07
/ PRIOR APPLICATION NUMBER: US 09/223,040
/ PRIOR FILING DATE: 1998-12-30
/ NUMBER OF SEQ ID NOS: 46
/ SOFTWARE: PatentIn Ver. 2.1
/ SEQ ID NO 44
/ LENGTH: 5
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence:flexible
/ OTHER INFORMATION: polylinker
US-09-287-849-44
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 4
US-09-147-142-31
/ Sequence 31, Application US/09147142
/ Patent No. US20020018749A1
/ GENERAL INFORMATION:
/ APPLICANT: HUDSON, Peter John
/ APPLICANT: KORTT, Alex Andrew
/ APPLICANT: IRVING, Robert Alexander
/ APPLICANT: ATWELL, John Leslie
/ TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
/ FILE REFERENCE: 016786/0212
/ CURRENT APPLICATION NUMBER: US/09/147,142
/ PRIOR FILING DATE: 1999-03-05
/ EARLIER APPLICATION NUMBER: PCT/AU98/00212
/ EARLIER FILING DATE: 1998-03-26
/ EARLIER APPLICATION NUMBER: AU PO 5917
/ EARLIER FILING DATE: 1997-03-27
/ NUMBER OF SEQ ID NOS: 32
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 31
/ LENGTH: 5
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Description of Artificial Sequence: peptide linker
US-09-147-142-31
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

```
RESULT 5
US-09-214-645-1
/ Sequence 1, Application US/09214645
/ Patent No. US2002028443A1
/ GENERAL INFORMATION:
/ APPLICANT: Short, Jay M.
/ TITLE OF INVENTION: METHOD OF DNA SHUFFLING WITH
/ TITLE OF INVENTION: POLYNUCLEOTIDES PRODUCED BY BLOCKING OR INTERRUPTING A
/ TITLE OF INVENTION: SYNTHESIS OR AMPLIFICATION PROCESS
/ FILE REFERENCE: DIVER1220-2 US/09/214,645
/ CURRENT APPLICATION NUMBER: US/09/214,645
/ PRIOR FILING DATE: 1999-09-27
/ PRIOR APPLICATION NUMBER: PCT/US97/12239
/ PRIOR FILING DATE: 1997-07-09
/ NUMBER OF SEQ ID NOS: 13
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 1
/ LENGTH: 5
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: linker peptide
US-09-214-645-1
```

```
Query Match          100.0%; Score 28; DB 9; Length 5;
Best Local Similarity 100.0%; Pred. No. 1.4e+06;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY      1 GGGGS 5
        |||||
Db       1 GGGGS 5
```

RESULT 6  
US-09-858-616-2  
; Sequence 2, Application US/09858616  
; Patent No. US20020031771A1  
; GENERAL INFORMATION:  
; APPLICANT: DIVERSA CORPORATION  
; APPLICANT: SHORT, Jay  
; TITLE OF INVENTION: SEQUENCE BASED SCREENING  
; FILE REFERENCE: DIVER1210-6  
; CURRENT APPLICATION NUMBER: US/09/858,616  
; PRIOR FILING DATE: 2001-09-10  
; PRIOR APPLICATION NUMBER: US 09/571,499  
; PRIOR FILING DATE: 2000-05-15  
; PRIOR APPLICATION NUMBER: US 09/557,276  
; PRIOR FILING DATE: 2000-04-24  
; PRIOR APPLICATION NUMBER: US 08/592,002  
; PRIOR FILING DATE: 1996-08-02  
; PRIOR APPLICATION NUMBER: US 60/008,317  
; PRIOR FILING DATE: 1995-12-07  
; PRIOR APPLICATION NUMBER: US 08/944,795  
; PRIOR FILING DATE: 1997-10-06  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 2  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial sequence  
; FEATURE:  
; OTHER INFORMATION: Linker peptide  
US-09-858-616-2

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5  
RESULT 7  
US-09-779-233-45  
; Sequence 45, Application US/09779233  
; Patent No. US20020045158A1  
; GENERAL INFORMATION:  
; APPLICANT: Casey, Casey  
; TITLE OF INVENTION: CELLS FOR DRUG DISCOVERY  
; FILE REFERENCE: 8325-0010  
; CURRENT APPLICATION NUMBER: US/09/779,233  
; CURRENT FILING DATE: 2001-02-08  
; NUMBER OF SEQ ID NOS: 45  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 45  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: linker  
US-09-779-233-45

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5  
RESULT 8  
US-09-989-789-3

; Sequence 3, Application US/09989789  
; Patent No. US20020063379A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Qiang  
; TITLE OF INVENTION: POSITION DEPENDENT RECOGNITION OF GNN NUCLEOTIDE  
; FILE REFERENCE: 8325-0011.20 / 511-US2  
; CURRENT APPLICATION NUMBER: US/09/989,789  
; CURRENT FILING DATE: 2002-03-25  
; NUMBER OF SEQ ID NOS: 4085  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Description of Artificial Sequence: peptide linker  
US-09-989-789-3

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 9  
US-09-976-787-21  
; Sequence 21, Application US/09976787  
; Patent No. US20020064528A1  
; GENERAL INFORMATION:  
; APPLICANT: Zhu, Zhenping  
; APPLICANT: Witte, Larry  
; TITLE OF INVENTION: Antibodies Specific to KDR and Uses Thereof  
; FILE REFERENCE: 11245/46505  
; CURRENT APPLICATION NUMBER: US/09/976,787  
; CURRENT FILING DATE: 2001-10-12  
; PRIOR APPLICATION NUMBER: US 09/493,539  
; PRIOR FILING DATE: 2000-01-28  
; PRIOR APPLICATION NUMBER: US 60/117,726  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 40  
; SOFTWARE: WordPerfect 8.0 for Windows  
; SEQ ID NO 21  
; LENGTH: 5  
; TYPE: PRT  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: peptide linker  
US-09-976-787-21

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GGGGS 5  
DB 1 GGGGS 5

RESULT 10  
US-09-192-854-180  
; Sequence 180, Application US/09192854  
; Patent No. US20020068276A1  
; GENERAL INFORMATION:  
; APPLICANT: Winter, Greg  
; APPLICANT: Tomlinson, Ian  
; TITLE OF INVENTION: Methods for Selecting Functional Peptides  
; FILE REFERENCE: 3789/72916  
; CURRENT APPLICATION NUMBER: US/09/192,854  
; CURRENT FILING DATE: 1998-11-17

EARLIER APPLICATION NUMBER: 60/066,729  
EARLIER FILING DATE: 1997-11-21  
NUMBER OF SEQ ID NOS: 212  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 180  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Homo sapiens  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Linker peptide  
US-09-192-854-180

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 11  
US-09-761-962-36  
Sequence 36, Application US/09761962  
Patent No. US2002007285A1  
GENERAL INFORMATION:  
APPLICANT: Memorial Sloan-Kettering Cancer Center  
TITLE OF INVENTION: Identification and Characterization of Multiple Splice  
TITLE OF INVENTION: Variants of Mu-  
TITLE OF INVENTION: Opioid Receptor (MOR-1) Gene  
FILE REFERENCE: 830002-2000.1  
CURRENT APPLICATION NUMBER: US/09/761,962  
PRIOR FILING DATE: 2001-01-17  
PRIOR APPLICATION NUMBER: 09/743,872  
PRIOR FILING DATE: 2001-03-13  
NUMBER OF SEQ ID NOS: 46  
SOFTWARE: PatentIn version 3.0  
SEQ ID NO 36  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: basic unit of a linking peptide  
US-09-761-962-36

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 12  
US-09-333-527-5  
Sequence 5, Application US/09333527  
Patent No. US20020078472A1  
GENERAL INFORMATION:  
APPLICANT: Paul CHRISTOU; Eva STROGER; Rainer FISCHER; Carmen MARTIN-VAQUERO; Stef  
TITLE OF INVENTION: METHODS AND MEANS FOR EXPRESSION OF MAMMALIAN POLYPEPTIDES  
NUMBER OF SEQUENCES: 43  
CORRESPONDENCE ADDRESS:  
ADDRESS: Fulbright & Jaworski L.L.P.  
STREET: 666 Fifth Avenue  
CITY: New York City  
STATE: New York  
COUNTRY: USA  
ZIP: 10103  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette, 3.25 inch, 1.44mb  
COMPUTER: IBM PS/2

OPERATING SYSTEM: PC-DOS  
SOFTWARE: Wordperfect  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/333,527  
FILING DATE: Concurrently Herewith  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 60/089,322  
FILING DATE: June 15, 1998  
ATTORNEY/AGENT INFORMATION:  
NAME: Mary Anne Schofield  
REGISTRATION NUMBER: 36,669  
REFERENCE/DOCKET NUMBER: KL/JIC 202.1 - JEL  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (212) 318-3000  
TELEFAX: (212) 752-5958  
INFORMATION FOR SEQ ID NO: 5:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 5  
TYPE: amino acid  
TOPOLOGY: linear  
US-09-333-527-5

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 13  
US-09-925-796-8  
Sequence 8, Application US/09925796  
Patent No. US20020081614A1  
GENERAL INFORMATION:  
APPLICANT: Case, Casey C.  
APPLICANT: Zhang, Lei  
TITLE OF INVENTION: Sangamo Biosciences, Inc.  
FILE REFERENCE: 019436-002000US  
CURRENT APPLICATION NUMBER: US/09/925,796  
PRIOR FILING DATE: 2001-08-09  
PRIOR APPLICATION NUMBER: 09/395,448  
PRIOR FILING DATE: 1999-09-14  
PRIOR APPLICATION NUMBER: 09/229,037  
PRIOR FILING DATE: 1999-01-12  
NUMBER OF SEQ ID NOS: 23  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 8  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: linker  
US-09-925-796-8

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
Db 1 GGGGS 5

RESULT 14  
US-09-815-837-116  
Sequence 116, Application US/09815837  
Patent No. US20020082411A1  
GENERAL INFORMATION:  
APPLICANT: Carter, Darrick

APPLICANT: Zhu, Shirley  
APPLICANT: Arimilli, Subhashini  
APPLICANT: Wang, Aijun  
APPLICANT: Corixa Corporation  
TITLE OF INVENTION: Immune Mediators and Related Methods  
FILE REFERENCE: 014058-005670US  
CURRENT APPLICATION NUMBER: US/09/815,837  
CURRENT FILING DATE: 2001-03-22  
PRIOR APPLICATION NUMBER: US 60/191,274  
PRIOR FILING DATE: 2000-03-22  
PRIOR APPLICATION NUMBER: US 60/204,249  
PRIOR FILING DATE: 2000-05-15  
PRIOR APPLICATION NUMBER: US 60/264,003  
PRIOR FILING DATE: 2001-01-23  
NUMBER OF SEQ ID NOS: 129  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 116  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence:downstream  
US-09-815-837-116

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

RESULT 15  
US-09-033-525-5  
Sequence 5, Application US/09033525  
Patent No. US20020090374A1  
GENERAL INFORMATION:  
APPLICANT: Yarkoni, Shai  
APPLICANT: Ben-Yehudah, Ahmi  
APPLICANT: Azar, Yehudith  
APPLICANT: Ageljan, Rami  
APPLICANT: Belotetsky, Ruth  
APPLICANT: Loterdoun-Galski, Haya  
TITLE OF INVENTION: CHIMERIC PROTEINS WITH CELL-TARGETING  
FILE REFERENCE: 9457-009-999  
CURRENT APPLICATION NUMBER: US/09/033,525  
CURRENT FILING DATE: 1998-03-02  
NUMBER OF SEQ ID NOS: 10  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 5  
LENGTH: 5  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Flexible polylinker  
US-09-033-525-5

Query Match 100.0%; Score 28; DB 9; Length 5;  
Best Local Similarity 100.0%; Pred. No. 1.4e+06;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
|||  
Db 1 GGGGS 5

Search completed: November 21, 2004, 14:03:48  
Job time : 3.18978 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 0.547445 Seconds  
(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	100.0	37	2 S29113	diptericin homolog
2	28	100.0	64	2 A86333	hypothetical prote
3	28	100.0	66	2 H84489	hypothetical prote
4	28	100.0	69	1 MIEC77	microcin B17 precu
5	28	100.0	78	2 E84686	hypothetical prote
6	28	100.0	80	2 T10550	hypothetical prote
7	28	100.0	81	2 PC2047	grain-sofiness pro
8	28	100.0	82	2 S19774	hypothetical prote
9	28	100.0	85	2 T32664	grain-sofiness pro
10	28	100.0	92	2 P00743	hypothetical prote
11	28	100.0	97	2 T48330	grain-sofiness pro
12	28	100.0	100	2 T49621	hypothetical prote
13	28	100.0	102	2 T25332	hypothetical prote
14	28	100.0	104	2 T02612	hypothetical prote
15	28	100.0	104	2 JC4190	holotricin 3 precu
16	28	100.0	108	2 G86252	hypothetical prote
17	28	100.0	109	2 S58673	RNA-binding protei
18	28	100.0	110	2 AC2391	RNA-binding protei
19	28	100.0	114	2 S28821	transcription fact
20	28	100.0	115	2 T35387	hypothetical prote
21	28	100.0	119	2 T07695	hypothetical prote
22	28	100.0	120	2 A81109	hypothetical prote
23	28	100.0	120	2 D83415	hypothetical prote
24	28	100.0	122	2 T04118	hypothetical prote
25	28	100.0	122	2 D86754	microchondrial proc
26	28	100.0	125	2 T16247	prophage p12 prote
27	28	100.0	128	2 T30428	hypothetical prote
28	28	100.0	131	2 H69062	molybdenum transpo
29	28	100.0	133	2 G75432	hypothetical prote

30	28	100.0	135	2 S55647	hypothetical prote
31	28	100.0	136	2 T02870	globulin 2 precurs
32	28	100.0	136	2 T29282	hypothetical prote
33	28	100.0	139	2 C87544	hypothetical prote
34	28	100.0	140	2 AC3088	hypothetical prote
35	28	100.0	144	2 S35716	glycine-rich prote
36	28	100.0	144	2 S04069	glycine-rich prote
37	28	100.0	144	2 T34730	probable gas vesic
38	28	100.0	145	1 UQ1062	glycine-rich prote
39	28	100.0	145	2 E84469	probable glycine-r
40	28	100.0	148	2 S46514	puroindoline-b pre
41	28	100.0	148	2 I38891	caudal-type homeot
42	28	100.0	149	2 T23179	hypothetical prote
43	28	100.0	150	2 C86224	hypothetical prote
44	28	100.0	152	2 T04811	stf1 protein homo
45	28	100.0	155	2 C86206	hypothetical prote

## ALIGNMENTS

RESULT 1  
S29113  
diptericin homolog - flesh fly (Sarcophaga peregrina)  
C/Species: Sarcophaga peregrina  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 09-Jul-2004  
C/Accession: S29113  
R/Ishikawa, M.; Kubo, T.; Natori, S.  
Biochem. J. 287, 573-578, 1992  
A/Title: Purification and characterization of a diptericin homologue from Sarcophaga per  
A/Reference number: S29113; MUID:93074996; PMID:1445217  
A/Accession: S29113  
A/Status: preliminary  
A/Molecule type: protein  
A/Residues: 1-37 <ISH>  
A/Cross-references: UNIPROT:Q9TWM2

Query Match 100.0%; Score 28; DB 2; Length 37;  
Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 19 GGGGS 23

## RESULT 2

A86333  
hypothetical protein T20H2.25 - Arabidopsis thaliana  
C/Species: Arabidopsis thaliana (mouse-ear cress)  
C/Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 09-Jul-2004  
C/Accession: A86333  
R/Theologis, A.; Becker, J.R.; Palm, C.J.; Federgraff, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasey, T.H.; Dewart, K.;  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 815-820, 2000  
A/Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Matli, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A/Authors: Salberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A/Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A/Reference number: A86141; MUID:21016719; PMID:11130712  
A/Accession: A86333  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-64 <STO>  
A/Cross-references: UNIPROT:Q9LNS5; GB:A8005172; NID:98779001; PIDN:AAF79916.1; GSPDB:GN  
C/Genetics:  
A/Map position: 1

Query Match 100.0%; Score 28; DB 2; Length 64;  
Best Local Similarity 100.0%; Pred. No. 2.2e+02;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 25 GGGGS 29

## RESULT 3

H84489  
 hypothetical protein At2g10020 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: H84489

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: H84489

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-66 <STO>

A/Cross-references: UNIPROT:Q9SK19; GB:AE002093; NID:g4556680; PIDN:AMD2697.1; GSPDB:GN

C/Genetics:

A/Gene: At2g10020

A/Map position: 2

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 66;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 22 GGGGS 26

## RESULT 4

MIEC77

microcin B17 precursor - Escherichia coli plasmid pmcCB17

C/Species: Escherichia coli

C/Date: 30-Jun-1988 #sequence\_revision 01-Dec-1995 #text\_change 09-Jul-2004

C/Accession: A25219; A32058; I41099; A58368; S67977

R/Davagnino, J.; Herrero, M.; Furlong, D.; Moreno, F.; Kolter, R.

Proteins 1, 230-238, 1986

A/Title: The DNA replication inhibitor microcin B17 is a forty-three-amino-acid protein

A/Reference number: A25219; MUID:88217867; PMID:3329725

A/Accession: A25219

A/Molecule type: DNA

A/Residues: 1-69 <DAV>

A/Cross-references: UNIPROT:P05834; GB:M15469; NID:g146787; PIDN:AAA24141.1; PID:g146786

J./Genlind, O.; Moreno, F.; Kolter, R.

A/Title: DNA sequence, products, and transcriptional pattern of the genes involved in pT

A/Reference number: A32058; MUID:89123111; PMID:2644225

A/Accession: A32058

A/Molecule type: DNA

A/Residues: 1-69 <GEN>

A/Cross-references: GB:M24253; NID:g311145; PIDN:AAA72741.1; PID:g522250

R./Connell, N.; Han, Z.; Moreno, F.; Kolter, R.

Mol. Microbiol. 1, 195-201, 1987

A/Title: An E. coli promoter induced by the cessation of growth.

A/Reference number: I41099; MUID:88216163; PMID:2835580

A/Accession: I41099

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-14 <CON>

A/Cross-references: EMBL:X06417; NID:g41978; PIDN:CAA29725.1; PID:g41979

R./Li, Y.M.; Milne, J.C.; Madison, L.L.; Kolter, R.; Welsh, C.T.

Science 274, 1188-1193, 1996

A/Title: From peptide precursors to oxazole and thiazole-containing peptide antibiotics:

A/Reference number: A58368; MUID:97053605; PMID:8895467

A/Accession: A58368

A/Molecule type: protein

A/Residues: 27-38 <LIY>

A/Experimental source: Escherichia coli strain ZK4(pY113)

A/Note: mass spectroscopy of peptides and biosynthetic intermediates

R./Jorgay, P.; Lee, U.; Koedel, U.; Vivas, E.; Warner, P.; Jedaratnam, D.; Kolter, R.

Proc. Natl. Acad. Sci. U.S.A. 91, 4519-4523, 1994

A/Title: Posttranslational modifications in microcin B17 define an additional class of DR

A/Reference number: A58375; MUID:94240167; PMID:8183941

A/Contents: annotation; (1)H-NMR spectroscopy of modified peptides

R./Bayer, A.; Freund, S.; Jung, G.

Eur. J. Biochem. 234, 414-426, 1995

A/Title: Post-translational heterocyclic backbone modifications in the 43-peptide antibiotic

A/Reference number: S67977; MUID:96128168; PMID:8536683

A/Accession: S67977

A/Status: preliminary

A/Molecule type: protein

A/Residues: 27-38 <BNV>

A/Genome: mcba

A/Genome: pmcCB17

A/Function: inhibits DNA gyrase, stopping DNA replication

A/Note: active against a large number of gram-negative enteric bacteria

C/Superfamily: microcin

C/Keywords: antibiotic; DNA replication inhibitor; oxazole/thiazole ring

F/1-26/Domain: signal sequence #status predicted <SIG>

F/27-69/Product: microcin B17 #status experimental <MNT>

F/39-40/Cross-link: oxazole (Gly-Ser) #status experimental

F/40-41/Cross-link: thiazole (Ser-Cys) #status experimental

F/47-48/Cross-link: thiazole (Gly-Cys) #status experimental

F/50-51/Cross-link: thiazole (Gly-Cys) #status experimental

F/55-56/Cross-link: thiazole (Gly-Cys) #status experimental

F/55-56/Cross-link: oxazole (Cys-Ser) #status experimental

F/61-62/Cross-link: oxazole (Gly-Ser) #status experimental

F/64-65/Cross-link: oxazole (Gly-Ser) #status experimental

Query Match

Best Local Similarity 100.0%; Score 28; DB 1; Length 69;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 36 GGGGS 40

## RESULT 5

E84686

hypothetical protein At2g28570 (imported) - Arabidopsis thaliana

C/Species: Arabidopsis thaliana (mouse-ear cress)

C/Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 09-Jul-2004

C/Accession: E84686

R/Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.

Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A/Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A/Reference number: A84420; MUID:20083487; PMID:10617197

A/Accession: E84686

A/Status: preliminary

A/Molecule type: DNA

A/Residues: 1-78 <STO>

A/Cross-references: UNIPROT:Q9SK01; GB:AE002093; NID:g4510404; PIDN:AMD21491.1; GSPDB:GN

C/Genetics:

A/Gene: At2g28570

A/Map position: 2

Query Match

Best Local Similarity 100.0%; Score 28; DB 2; Length 78;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
 |||||  
 Db 53 GGGGS 57

RESULT 6  
T10550  
hypothetical protein T12G13.70 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 16-Jul-1999 #sequence\_revision 16-Jul-1999 #text\_change 09-Jul-2004  
C:Accession: T10550  
R:Bevan, M.; Lennard, N.; Quail, M.; Harris, B.; Rajandream, M.A.; Barrell, B.G.; Bancro  
submitted to the Protein Sequence Database, June 1999  
A:Reference number: Z16533  
A:Accession: T10550  
A:Molecule type: DNA  
A:Residues: 1-80 <BEV>  
A:Cross-references: UNIPROT:Q9SUF7; EMBL:AF090252; GSPDB:GN00062; ATSP:T12G13.70  
C:Genetics:  
A:Gene: ATSP:T12G13.70  
A:Map position: 4

Query Match 100.0%; Score 28; DB 2; Length 80;  
Best Local Similarity 100.0%; Pred. No. 2.6e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 27 GGGGS 31

RESULT 7  
PC2047  
grain-softness protein - wheat (fragments)  
C:Species: Triticum aestivum (common wheat)  
C:Date: 05-Aug-1994 #sequence\_revision 05-Aug-1994 #text\_change 09-Jul-2004  
C:Accession: PC2047  
R:Jolly, C.J.; Rahman, S.; Kortt, A.A.; Higgins, T.J.V.  
Theor. Appl. Genet. 86, 589-597, 1993  
A:Title: Characterisation of the wheat Mr 15000 grain-softness protein and analysis of t  
A:Reference number: P00743  
A:Accession: PC2047  
A:Molecule type: protein  
A:Residues: 1-18;19-24;25-31;32-38;39-45;46-51;52-56;57-60;61-65;66-71;72-77;78-81 <JOL>  
A:Cross-references: UNIPROT:Q7M1T8  
C:Comment: This protein is the product of the Ha locus and thus be the major factor that  
C:Keywords: seed

Query Match 100.0%; Score 28; DB 2; Length 81;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 4 GGGGS 8

RESULT 8  
S19774  
glycine-rich protein - tomato (fragment)  
C:Species: Lycopersicon esculentum (tomato)  
C:Date: 30-Jun-1992 #sequence\_revision 30-Jun-1992 #text\_change 09-Jul-2004  
C:Accession: S19774  
R:Parsons, B.L.  
submitted to the EMBL Data Library, May 1991  
A:Reference number: S19773  
A:Accession: S19774  
A:Molecule type: mRNA  
A:Residues: 1-82 <PAR>  
A:Cross-references: UNIPROT:Q04130; EMBL:X59883; NID:g19321; PIDN:CAA42538.1; PID:g19322  
C:Superfamily: glycine-rich RNA-binding protein; ribonucleoprotein repeat homology

Query Match 100.0%; Score 28; DB 2; Length 82;  
Best Local Similarity 100.0%; Pred. No. 2.7e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 72 GGGGS 76

RESULT 9  
T32664  
hypothetical protein F16B4.7 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T32664  
R:Davidson, S.; Wohlmann, P.; Bauer, C.; O'Neal, D.  
submitted to the EMBL Data Library, December 1997  
A:Description: The sequence of C. elegans cosmid F16B4.  
A:Reference number: Z21208  
A:Accession: T32664  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-85 <DAV>  
A:Cross-references: UNIPROT:Q44629; EMBL:AF039048; PIDN:AAE94238.1; GSPDB:GN00023; CESP:  
A:Experimental source: strain Bristol N2; clone F16B4  
C:Genetics:  
A:Gene: CESP:F16B4.7  
A:Map position: 5  
A:introns: 36/1  
C:Superfamily: Arabidopsis glycine-rich protein 3

Query Match 100.0%; Score 28; DB 2; Length 85;  
Best Local Similarity 100.0%; Pred. No. 2.8e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 39 GGGGS 43

RESULT 10  
P00743  
grain-softness protein - wheat (fragments)  
C:Species: Triticum aestivum (common wheat)  
C:Date: 19-May-1994 #sequence\_revision 19-May-1994 #text\_change 09-Jul-2004  
C:Accession: P00743  
R:Jolly, C.J.; Rahman, S.; Kortt, A.A.; Higgins, T.J.V.  
Theor. Appl. Genet. 86, 589-597, 1993  
A:Title: Characterisation of the wheat Mr 15000 grain-softness protein and analysis of t  
A:Reference number: P00743  
A:Accession: P00743  
A:Molecule type: protein  
A:Residues: 1-92 <JOL>  
A:Cross-references: UNIPROT:Q7M1T7  
A:Experimental source: seed  
C:Keywords: seed

Query Match 100.0%; Score 28; DB 2; Length 92;  
Best Local Similarity 100.0%; Pred. No. 3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
|||||  
Db 4 GGGGS 8

RESULT 11  
T48330  
hypothetical protein F15A17.120 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 09-Jul-2004  
C:Accession: T48330  
R:Bevan, M.; Terry, N.; Ardiles, W.; Buyschaert, C.; Dasseville, R.; De Clerck, R.; De  
ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
submitted to the Protein Sequence Database, April 2000  
A:Reference number: Z24491

A;Accession: T48330  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-97 <BEV>  
A;Cross-references: UNIPROT:Q9LYX6; EMBL:AL163002  
A;Experimental source: cultivar Columbia; BAC clone F15A17  
C;Genetics:  
A;Map position: 5  
A;Intons: 7/1  
A;Note: F15A17.120

Query Match 100.0%; Score 28; DB 2; Length 97;  
Best Local Similarity 100.0%; Pred. No. 3.1e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 47 GGGGS 51

## RESULT 12

T49621  
hypothetical protein B5022.30 [imported] - Neurospora crassa  
C;Species: Neurospora crassa  
C;Date: 02-Jun-2000 #sequence\_revision 02-Jun-2000 #text\_change 02-Jun-2000  
C;Accession: T49621  
R;Schulte, U.; Alyn, V.; Hohseil, J.; Brandt, P.; Fartmann, B.; Holland, R.; Nyakatura,  
submitted to the Protein Sequence Database, May 2000  
A;Reference number: Z25022  
A;Accession: T49621  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-100 <SCH>  
A;Cross-references: EMBL:AJ355932; GSPDB:GN00116; NCSP:B5022.30  
A;Experimental source: BAC clone B5022; strain OR74A  
C;Genetics:  
A;Gene: NCSP:B5022.30  
A;Map position: 6  
A;Intons: 22/1; 52/1

Query Match 100.0%; Score 28; DB 2; Length 100;  
Best Local Similarity 100.0%; Pred. No. 3.2e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 11 GGGGS 15

## RESULT 13

T25332  
hypothetical protein T26H5.4 - Caenorhabditis elegans  
C;Species: Caenorhabditis elegans  
C;Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C;Accession: T25332  
R;Gardner, A.  
submitted to the EMBL Data Library, November 1996  
A;Reference number: Z20017  
A;Accession: T25332  
A;Status: preliminary; translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-102 <WIL>  
A;Cross-references: UNIPROT:O62385; EMBL:Z82056; PIDN:CA804855.1; GSPDB:GN00023; CESP:T2  
C;Experimental source: clone T26H5  
C;Genetics:  
A;Gene: CESP:T26H5.4  
A;Map position: 5  
A;Intons: 13/1; 96/1  
C;Superfamily: hypothetical protein K01D12.8

Query Match 100.0%; Score 28; DB 2; Length 102;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 52 GGGGS 56

## RESULT 14

T02612  
hypothetical protein At2g26120 [imported] - Arabidopsis thaliana  
N;Alternate names: hypothetical protein T19L18.7  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 24-Mar-1999 #sequence\_revision 24-Mar-1999 #text\_change 09-Jul-2004  
C;Accession: T02612; F84656  
R;Rounsley, S.D.; Kaul, S.; Lin, X.; Kechum, K.A.; Crosby, M.L.; Brandon, R.C.; Sykes, S.  
submitted to the EMBL Data Library, August 1998  
A;Description: Arabidopsis thaliana chromosome II BAC T19L18 genomic sequence.  
A;Reference number: Z14681  
A;Accession: T02612  
A;Status: translated from GB/EMBL/DBJ  
A;Molecule type: DNA  
A;Residues: 1-104 <ROU>  
A;Cross-references: UNIPROT:O80985; EMBL:AC004747; NID:G3413696; PID:G3413702  
A;Experimental source: cultivar Columbia  
R;Lin, X.; Kaul, S.; Rounsley, S.D.; Shee, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.  
euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J  
Nature 402, 761-768, 1999  
A;Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A;Reference number: A84420; MUID:20083487; PMID:10617197  
A;Accession: F84656  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-104 <STO>  
A;Cross-references: GB:AE002093; NID:G3413702; PIDN:AA031225.1; GSPDB:GN00139  
C;Genetics:  
A;Gene: T19L18.7; At2g26120  
A;Map position: 2  
A;Intons: 49/3

Query Match 100.0%; Score 28; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5  
Db 96 GGGGS 100

## RESULT 15

JC4190  
holotricin 3 precursor - Holotrichia diomphalia  
N;Alternate names: antifungal protein  
C;Species: Holotrichia diomphalia  
C;Date: 04-Oct-1995 #sequence\_revision 10-Nov-1995 #text\_change 09-Jul-2004  
C;Accession: JC4190  
R;Lee, S.Y.; Moon, H.J.; Kurata, S.; Natori, S.; Lee, B.L.  
Biol. Pharm. Bull. 18, 1049-1052, 1995  
A;Title: Purification and cDNA cloning of an antifungal protein from the hemolymph of Ho  
A;Reference number: JC4190; MUID:96073722; PMID:8535353  
A;Accession: JC4190  
A;Molecule type: mRNA  
A;Residues: 1-104 <LEE>  
A;Cross-references: UNIPROT:Q25055; DDBJ:DJ3744; NID:G1088433; PIDN:BA02889.1; PID:d100;  
C;Comment: This protein is a Gly- and His-rich protein and a constitutive protein of lar  
C;Keywords: hemolymph  
F;1-20/Domain: signal sequence #status predicted <SIG>  
F;21-104/Product: holotricin 3 #status predicted <MAT>

Query Match 100.0%; Score 28; DB 2; Length 104;  
Best Local Similarity 100.0%; Pred. No. 3.3e+02;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GGGGS 5

|||||  
Db 69 GGGGS 73

Search completed: November 21, 2004, 13:43:26  
Job time : 2.54745 secs

---

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:28:31 ; Search time 2.77372 Seconds  
(without alignments)  
1037.189 Million cell updates/sec

Title: US-09-936-702-1  
Perfect score: 28  
Sequence: 1 GGGGS 5

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues  
Total number of hits satisfying chosen parameters: 1825181

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Uniprot 02: \*  
1: uniprot\_aprot: \*  
2: uniprot\_trembl: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	28	100.0	10	06X2S9	06X2S9 homo sapien
2	28	100.0	10	AAP83315	AAP83315 homo sapi
3	28	100.0	17	PKX5_PRRAM	P82617 periplaneta
4	28	100.0	17	09R4Y9	09R4Y9 alcailligenes
5	28	100.0	17	064450	064450 mus musculus
6	28	100.0	20	08W238	08W238 glycine max
7	28	100.0	20	09R582	09R582 neisseria m
8	28	100.0	23	09UC00	09UC00 homo sapien
9	28	100.0	26	042226	042226 arabidopsis
10	28	100.0	31	09F115	09F115 enterococcus
11	28	100.0	37	09TWM2	09TWM2 sarcophaga
12	28	100.0	40	013832	013832 homo sapien
13	28	100.0	40	013833	013833 homo sapien
14	28	100.0	44	06Y1D5	06Y1D5 lactuca sat
15	28	100.0	44	AAR22556	AAR22556 lactuca s
16	28	100.0	45	07XMY7	07XMY7 oryza sativ
17	28	100.0	48	07QSB6	07QSB6 giardia lam
18	28	100.0	50	084Y55	084Y55 oryza sativ
19	28	100.0	50	0943L9	0943L9 oryza sativ
20	28	100.0	50	096C82	096C82 rhizobium l
21	28	100.0	53	07S2G8	07S2G8 neuropeptid
22	28	100.0	53	08LNM3	08LNM3 oryza sativ
23	28	100.0	54	06YV44	06YV44 oryza sativ
24	28	100.0	54	BAD17456	BAD17456 oryza sat
25	28	100.0	55	0751X3	0751X3 oryza sativ
26	28	100.0	55	AAS07084	AAS07084 oryza sat
27	28	100.0	60	08H647	08H647 oryza sativ
28	28	100.0	61	082GM4	082GM4 streptomyces
29	28	100.0	62	08N2Z2	08N2Z2 homo sapien
30	28	100.0	62	07B2W2	07B2W2 oryza sativ
31	28	100.0	62	BAD01701	BAD01701 oryza sat

32	28	100.0	62	2	BAD03150	BAD03150 oryza sat
33	28	100.0	63	2	094E89	094E89 oryza sativ
34	28	100.0	63	2	08WB65	08WB65 white spot
35	28	100.0	64	2	06K665	06K665 oryza sativ
36	28	100.0	64	2	09LNS5	09LNS5 arabidopsis
37	28	100.0	64	2	06ERC8	06ERC8 oryza sativ
38	28	100.0	65	2	06Z5J7	06Z5J7 oryza sativ
39	28	100.0	65	2	094J57	094J57 oryza sativ
40	28	100.0	65	2	BAD17269	BAD17269 oryza sat
41	28	100.0	66	2	09XZQ8	09XZQ8 caenorhabdi
42	28	100.0	66	2	09XT04	09XT04 macropus ru
43	28	100.0	66	2	08H3Q4	08H3Q4 oryza sativ
44	28	100.0	66	2	09SK19	09SK19 arabidopsis
45	28	100.0	67	2	06YVW2	06YVW2 oryza sativ

## ALIGNMENTS

```

RESULT 1
ID 06X2S9 PRELIMINARY; PRT; 10 AA.
AC 06X2S9;
DT 05-JUL-2004 (TREMBLrel. 27, Created)
DT 05-JUL-2004 (TREMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE Translocase of the inner mitochondrial membrane (Fragment).
GN Name=TIM23;
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX MEDLINE=22890074; PubMed=14527731;
RA Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.;
RT "Human poly(ADP-ribose) glycohydrolase (PARC) gene and the common
RT translocase 23 (TIM23).";
RL Gene 314:181-190(2003).
DR EMBL, AY258588; AAP83315.1; -.
FT NON_TER
SQ
SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;
Query Match 100.0%; Score 28; DB 2; Length 10;
Best Local Similarity 100.0%; Pred. No. 2.2e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5
   |||||
Db 3 GGGGS 7

RESULT 2
ID AAP83315 PRELIMINARY; PRT; 10 AA.
AC AAP83315;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE Translocase of the inner mitochondrial membrane (Fragment).
GN TIM23.
OS Homo sapiens (Human);
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Skin;
RX Meyer R.G., Meyer-Ficca M.L., Jacobson E.L., Jacobson M.K.;
RT "Human poly(ADP-ribose) glycohydrolase (PARC) gene and the common
RT promoter sequence it shares with inner mitochondrial membrane

```

RT translocase 23 (TIM23).";  
 RL Gene 314:181-190(2003).  
 DR EMBL: AY258588; AAP83315.1; -.  
 FT NON\_TER 10 10  
 SQ SEQUENCE 10 AA; 893 MW; 2D7A7165B8787878 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 10;  
 Best Local Similarity 100.0%; Pred. No. 2.2e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 3 GGGGS 7

## RESULT 3

PKS\_PERAM STANDARD; PRT; 17 AA.  
 ID PK2617;  
 AC 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 05-JUN-2004 (Rel. 44, Last annotation update)  
 DE Pyrokinin-5 (Pea-PK-5) (FKPRL-amide).  
 OS Periplaneta americana (American cockroach).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Orthopteroidea; Dictyoptera; Blattaria; Blattellidae;  
 OC Blattellidae; Periplaneta.  
 NC NCBI\_TaxID=6978;

SEQUENCE, FUNCTION, AND MASS SPECTROMETRY.  
 RP TISSUE=Abdominal perisymphathetic organs;  
 RC MEDLINE=99212469; PubMed=10196736;  
 RA Predel R., Kellner R., Nachman R.J., Holman G.M., Rapus J., Gaede G.;  
 RT "Differential distribution of pyrokinin-isoforms in cerebral and  
 abdominal neuromal organs of the American cockroach.";  
 RL Insect Biochem. Mol. Biol. 29:139-144(1999).  
 RN [2]

TISSUE SPECIFICITY.  
 RP MEDLINE=20189894; PubMed=10723010;  
 RA Predel R., Eckert M.;  
 RT "Tissue-specific distribution of FKPRamide in the nervous system of  
 the American cockroach.";  
 RL J. Comp. Neurol. 419:352-363(2000).  
 CC -1- FUNCTION: Mediates visceral muscle contractile activity (myotropic  
 activity).  
 CC -1- TISSUE SPECIFICITY: Mainly in abdominal perisymphathetic organs and  
 to a lesser extent in retrocerebral complex.  
 CC -1- MASS SPECTROMETRY: MW=1651.7; METHOD=MALDI; RANGE=1-17;  
 CC NOTES=Ref.1.  
 CC -1- SIMILARITY: Belongs to the pyrokinin family.  
 DR InterPro: IPR001484; Pyrokinin.  
 DR PROSITE: PS00539; PYROKININ; 1.  
 KW Annotation: Direct protein sequencing; Neuropeptide; Pyrokinin.  
 FT MOD\_RES 17  
 LEUCINE amide.  
 SQ SEQUENCE 17 AA; 1653 MW; 8527162BA45BA54 CRC64;

Query Match 100.0%; Score 28; DB 1; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 1 GGGGS 5

## RESULT 4

O9R4Y9 PRELIMINARY; PRT; 17 AA.  
 AC O9R4Y9;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE Aromatic amine dehydrogenase beta subunit (Fragment).

OS Alcaligenes faecalis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;  
 OC Alcaligenaceae; Alcaligenes.  
 NC NCBI\_TaxID=511;  
 RN [1]

SEQUENCE.  
 RX MEDLINE=94245619; PubMed=8188594;  
 RA Govindaraj S., Eisenstein E., Jones L.H., Sanders-Loehr J.,  
 RA Chiserosov A.Y., Davidson V.L., Edwards S.L.;  
 RT "Aromatic amine dehydrogenase, a second tryptophan tryptophylquinone  
 enzyme.";  
 RL J. Bacteriol. 176:2922-2929(1994).  
 SQ SEQUENCE 17 AA; 1510 MW; 6EEBAEB9D89D2661 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 2 GGGGS 6

## RESULT 5

O64450 PRELIMINARY; PRT; 17 AA.  
 ID O64450;  
 AC 064450;  
 DT 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Uridine kinase (EC 2.7.1.48) (Fragment).  
 GN Name=Umpk;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NC NCBI\_TaxID=10090;  
 RN [1]

SEQUENCE FROM N.A.  
 RP TISSUE=Liver;  
 RC ROPP P.A., Traut T.W.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: U57332; AAB01998.1; -.  
 DR MGI: 98904; Umpk.  
 DR GO: GO:0016301; P:kinase activity; IEA.  
 DR GO: GO:0016740; F:transferase activity; IEA.  
 DR GO: GO:0004849; F:uridine kinase activity; IEA.  
 KW Kinase; Transferase.  
 FT NON\_TER 17  
 SQ SEQUENCE 17 AA; 1464 MW; 14E427CBA1168634 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 17;  
 Best Local Similarity 100.0%; Pred. No. 3.7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 5 GGGGS 9

## RESULT 6

O8W238 PRELIMINARY; PRT; 20 AA.  
 ID O8W238;  
 AC O8W238;  
 DT 01-MAR-2002 (TREMBlrel. 20, Created)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last sequence update)  
 DT 01-MAR-2002 (TREMBlrel. 20, Last annotation update)  
 DE GR-2 factor (Fragment).  
 OS Glycine max (Soybean).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliopsida; core eudicots; rosids;  
 OC euroids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae; Glycine.  
 NC NCBI\_TaxID=3847;  
 RN [1]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=21471140; PubMed=11587508;  
 RA O'Grady K., Gokhjian V.H., Naim C.J., Nagao R.T., Key J.L.;  
 RT "The transcript abundance of GmGT-2, a new member of the GT-2 family  
 of transcription factors from soybean, is down-regulated by light in a  
 RT phytochrome-dependent manner.";  
 RL Plant Mol. Biol. 47:367-378(2001).  
 DR EMBL: AF372500; AAL65126.1; -;  
 FT NON TER 20 20  
 SQ SEQUENCE 20 AA; 1692 MW; F65C75CD9C6B63B CRC64;

Query Match 100.0%; Score 28; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 7  
 Q9R582 PRELIMINARY; PRT; 20 AA.  
 AC Q9R582;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Transferrin-binding protein 2 (Fragment).  
 OS Neisseria meningitidis.  
 OC Bacteria; Proteobacteria; Betaproteobacteria; Neisseriales;  
 OC Neisseriaceae; Neisseria.  
 OX NCBI\_TaxID=487;  
 RN [1]  
 RP SEQUENCE.  
 RA MEDLINE=93307625; PubMed=8319886;  
 RA Griffiths E., Stevenson P., Byfield P., Al-Alderaan D.A.,  
 RA Borriello S.P., Holland J., Parsons T., Williams P.;  
 RT "Antigenic relationships of transferrin-binding proteins from  
 RT Neisseria meningitidis, N. gonorrhoeae and Haemophilus influenzae:  
 RT cross-reactivity of antibodies to NH2-terminal peptides.";  
 RL FEMS Microbiol. Lett. 109:85-91(1993).  
 SQ SEQUENCE 20 AA; 1977 MW; 6000EB169F09227E CRC64;

Query Match 100.0%; Score 28; DB 2; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 4.3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 3 GGGGS 7

RESULT 8  
 Q9UC00 PRELIMINARY; PRT; 23 AA.  
 AC Q9UC00;  
 DT 01-MAY-2000 (TRENBLrel. 13, Created)  
 DT 01-MAY-2000 (TRENBLrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TRENBLrel. 24, Last annotation update)  
 DE Enhancement of wound HEALING process.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=95130623; PubMed=7829572;  
 RA Pletschbacher M.D., Polarek J.W., Craig W.S., Tschopp J.F.,  
 RA Sipes N.J., Harper J.R.;  
 RL J. Cell. Biochem. 56:150-154(1994).  
 DR GO: 0009611; P:response to wounding; TAS.  
 SQ SEQUENCE 23 AA; 2268 MW; CE7399CB9903891 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 23;  
 Best Local Similarity 100.0%; Pred. No. 4.9e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 9  
 Q42226 PRELIMINARY; PRT; 26 AA.  
 AC Q42226;  
 DT 01-NOV-1996 (TRENBLrel. 01, Created)  
 DT 01-NOV-1996 (TRENBLrel. 01, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Seed maturation protein (Fragment).  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Dry seeds of A. thaliana ecotype Columbia;  
 RA CNRS;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Dry seeds of A. thaliana ecotype Columbia;  
 RA Raynal M., Grellet F., Lande M., Meyer Y., Cooke R., Delserny M.;  
 RL Submitted (FEB-1994) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: Z29850; CAA82818.1; -;  
 FT NON TER 1 1  
 SQ SEQUENCE 26 AA; 2370 MW; 6E0902E39464466A CRC64;

Query Match 100.0%; Score 28; DB 2; Length 26;  
 Best Local Similarity 100.0%; Pred. No. 5.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 DB 11 GGGGS 15

RESULT 10  
 Q9F115 PRELIMINARY; PRT; 31 AA.  
 AC Q9F115;  
 DT 01-MAR-2001 (TRENBLrel. 16, Created)  
 DT 01-MAR-2001 (TRENBLrel. 16, Last sequence update)  
 DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)  
 DE Hypothetical protein EP0010.  
 GN Name=EP0010;  
 OS Enterococcus faecalis (Streptococcus faecalis).  
 OC Bacteria; Firmicutes; Lactobacillales; Enterococcaceae; Enterococcus.  
 OX NCBI\_TaxID=1351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=20453452; PubMed=10998166;  
 RA De Boever E.H., Clewell D.B., Fraser C.M.;  
 RT "Enterococcus faecalis conjugative plasmid pAM373: complete nucleotide  
 RT sequence and genetic analyses of sex pheromone response.";  
 RL Mol. Microbiol. 37:1327-1341(2000).  
 DR EMBL: AR002565; AAG40421.1; -;  
 KW Hypothetical protein; Plasmid.  
 SQ SEQUENCE 31 AA; 3509 MW; 4E19CB94B3DB9421 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 6.5e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 18 GGGGS 22

## RESULT 11

Q9TWM2 PRELIMINARY; PRT; 37 AA.  
 ID 09TWM2  
 AC 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE Dipterician homolig (Fragment).  
 OS Sarcophaga peregrina (Flesh fly) (Boettcherisca peregrina).  
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Cecidoidea;  
 OC Sarcophagidae; Sarcophaga.  
 NCBI\_TaxID=7386;  
 OX NCBI [1]  
 RP SEQUENCE.  
 RC TISSUE=LARVAL HEMOLYMPH;  
 RX MEDLINE=93074996; PubMed=1445217;  
 RA Ishikawa M., Kudo T., Natori S.;  
 RT "Purification and characterization of a dipterician homologue from  
 RT Sarcophaga peregrina (Flesh fly).";  
 RL Biochem. J. 287:573-578(1992).  
 CC -1- FUNCTION: BACTERICIDAL ACTIVITY AGAINST GRAM-NEGATIVE BACTERIA  
 CC E. COLI AND S. SONNET.  
 CC -1- TISSUE SPECIFICITY: SYNTHESIZED BY THE FAT BODY AND IS SECRETED  
 CC INTO THE HEMOLYMPH.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSION IN THE LARVAE STARTS A FEW MINUTES  
 CC AFTER THE INJURY OF THE BODY WALL REACHING A MAXIMUM AFTER ABOUT  
 CC 10 HOURS. THE MAXIMUM LASTS FOR AT LEAST 3 HOURS.  
 CC -1- INDUCTION: IN RESPONSE TO INJURY OF THE BODY WALL OF THE LARVAE.  
 DR PIR; S29113; S29113.  
 DR GO; GO:0006952; P:defense response; IEA.  
 DR GO; GO:0006805; P:xenobiotic metabolism; IEA.  
 KM Antibiotic; Insect immunity.  
 FT DOMAIN 18 22 POLY-GLY.  
 FT NON\_TER 37 37  
 SQ SEQUENCE 37 AA; 3928 MW; E3BAC8105D2DABC7 CRC64;  
 Query Match 100.0%; Score 28; DB 2; Length 37;  
 Best Local Similarity 100.0%; Pred. No. 7; 7e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGGS 5  
 |||||  
 Db 19 GGGGS 23

RESULT 12  
 Q13832 PRELIMINARY; PRT; 40 AA.  
 ID 013832  
 AC 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last annotation update)  
 DE B2 bradykinin receptor basal promoter, allele BP-58-C (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 OX NCBI [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96209920; PubMed=8655154;  
 RA Braun A., Maier E., Kammerer S., Mueller B., Roscher A.A.;  
 RT "A novel sequence polymorphism in the promoter region of the human  
 RT bradykinin B2-receptor gene.";  
 RL Hum. Genet. 97:688-689(1996).  
 DR EMBL; X91663; CAA62851.1; -;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 KM Receptor.

FT NON\_TER 1 1  
 FT NON\_TER 40 40  
 SQ SEQUENCE 40 AA; 4140 MW; 3908E9AD371EF4A5 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 8; 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 6 GGGGS 10

## RESULT 13

Q13833 PRELIMINARY; PRT; 40 AA.  
 ID 013833  
 AC 01-NOV-1996 (TREMBlrel. 01, Created)  
 DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
 DT 01-JUN-2003 (TREMBlrel. 24, Last annotation update)  
 DE B2 bradykinin receptor basal promoter, allele BP-58-T (Fragment).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 OX NCBI [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=96209920; PubMed=8655154;  
 RA Braun A., Maier E., Kammerer S., Mueller B., Roscher A.A.;  
 RT "A novel sequence polymorphism in the promoter region of the human  
 RT bradykinin B2-receptor gene.";  
 RL Hum. Genet. 97:688-689(1996).  
 DR EMBL; X91664; CAA62852.1; -;  
 DR GO; GO:0004872; F:receptor activity; IEA.  
 KM Receptor.  
 FT NON\_TER 1 1  
 FT NON\_TER 40 40  
 SQ SEQUENCE 40 AA; 4153 MW; 1408E9AD371EF17F CRC64;  
 Query Match 100.0%; Score 28; DB 2; Length 40;  
 Best Local Similarity 100.0%; Pred. No. 8; 3e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGGS 5  
 |||||  
 Db 6 GGGGS 10

RESULT 14  
 Q6Y1D5 PRELIMINARY; PRT; 44 AA.  
 ID 06Y1D5  
 AC 05-JUL-2004 (TREMBlrel. 27, Created)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
 DE 60S acidic ribosomal protein P3 (Fragment).  
 OS Lactuca sativa (Garden lettuce).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asteride;  
 OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
 OC Lactuca.  
 NCBI\_TaxID=4236;  
 OX NCBI [1]  
 RP SEQUENCE FROM N.A.  
 RX Kuang H., Nevo E., Michelmore R.W.;  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ database.  
 DR EMBL; AY93418; AAR22556.1; -;  
 DR GO; GO:0003735; F:structural constituent of ribosome; IEA.  
 KM Ribosomal protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 44 44  
 SQ SEQUENCE 44 AA; 4358 MW; A9FF3A8D93744CD9 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 39 GGGGS 43

## RESULT 15

AA022556 PRELIMINARY; PRT; 44 AA.  
 ID AA022556;  
 AC AA022556;  
 DT 02-MAR-2004 (TRENBLREL. 27, Created)  
 DT 02-MAR-2004 (TRENBLREL. 27, Last sequence update)  
 DT 02-MAR-2004 (TRENBLREL. 27, Last annotation update)  
 DS 60S acidic ribosomal protein P3 (Fragment).  
 OS Lactuca sativa (Garden lettuce).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;  
 OC campanulids; Asterales; Asteraceae; Cichorioideae; Cichorieae;  
 OC Lactuca.  
 OK NCBI\_TaxID=4236;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=cv. Calmar;  
 RA Kuang H., Nevo E., Michelmore R.W.;  
 RT "Heterogeneous Rates of Evolution at the Major Cluster Disease  
 Resistance Genes in Lettuce due to Distinct Genetic Processes."  
 RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AY193418; AA022556.1; -.  
 KW Ribosomal protein.  
 FT NON\_TER 1 1  
 FT 44 44  
 SQ SEQUENCE 44 AA; 4358 MW; A9F3A8D93744CD9 CRC64;

Query Match 100.0%; Score 28; DB 2; Length 44;  
 Best Local Similarity 100.0%; Pred. No. 9.1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGS 5  
 |||||  
 Db 39 GGGGS 43

Search completed: November 21, 2004, 13:42:21  
 Job time : 5.77372 secs

**THIS PAGE BLANK (USPTO)**



CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure

XX  
 SQ Sequence 35 AA;

Query Match 100.0%; Score 196; DB 3; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 6, 6e-12;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35  
 DB 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35

# RESULT 2

ABJ39015 ID ABJ39015 standard; protein; 35 AA.

XX AC ABJ39015;

DT 09-OCT-2003 (first entry)

DE Synthetic linker Seq ID30 amino acid sequence.

XX KM Streptavidin fusion protein; SA; Streptomyces avidinil; biotin; scFvSA;  
 KM anti-CD25 antibody; antigen binding fragment; cytosolic; gene therapy;  
 KM cell-specific targeting agents; single chain; tumour.

XX OS Synthetic.

XX PN WO2003050260-A2.

XX PD 19-JUN-2003.

XX PF 06-DEC-2002; 2002WO-US039429.

XX PR 07-DEC-2001; 2001US-00013173.

XX PR 17-MAY-2002; 2002US-00150762.

XX PR 16-SEP-2002; 2002US-00244621.

XX PA (NEOR-) NEORX CORP.

XX PI Goshorn SC, Graves SS, Schultz JE, Lin Y, Sanderson JA, Reno JM;

XX PI Dearstyne EA;

XX PT WPI; 2003-532908/50.

XX PS New vector construct for expressing genomic streptavidin fusion proteins  
 XX useful as diagnostic markers or as cell-specific targeting agents.

XX PS Example 4; Page 51; 15pp; English.

CC This invention relates to novel vector constructs for the expression of  
 CC streptavidin fusion proteins. Streptavidin (SA) is produced by  
 CC Streptomyces avidinil and specifically binds water-soluble biotin. The  
 CC vectors comprise a nucleic acid sequence encoding genomic streptavidin, a  
 CC promoter operatively linked to the nucleic acid sequence and a cloning  
 CC site for insertion of a second nucleic acid sequence encoding an anti-  
 CC CD25 antibody or its antigen binding fragment. The fusion proteins  
 CC encoded are known as single chain antibody-genomic streptavidin fusion  
 CC proteins (scFvSA). The vectors may have cytosolic activity when used in  
 CC gene therapy. The vectors may be useful in expressing genomic

CC streptavidin fusion cassettes. The fusion proteins may be used as  
 CC diagnostic markers or as cell-specific targeting agents. These may also  
 CC be used in treating tumours. The present sequence is that of synthetic  
 CC amino acid linker Seq ID30 used in the streptavidin fusion proteins of  
 CC the invention

XX SQ Sequence 35 AA;

Query Match 100.0%; Score 196; DB 6; Length 35;  
 Best Local Similarity 100.0%; Pred. No. 6, 6e-12;  
 Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35  
 DB 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35

# RESULT 3

AAV87573 ID AAV87573 standard; protein; 40 AA.

XX AC AAV87573;

DT 18-JUL-2000 (first entry)

DE Linker between CH3 and FcL domains of Ig-FcL fusion protein.

XX KM Linker; immunoglobulin; Ig; CH3 domain; Fas ligand; FasL; fusion protein;  
 KM targeted apoptosis; cell surface binding protein; specific cytotoxicity;  
 KM autoimmune disorder; inflammatory condition; cancer.

XX OS Synthetic.

XX PN US6046310-A.

XX PD 04-APR-2000.

XX PF 11-MAR-1997; 97US-00815190.

XX PR 13-MAR-1996; 96US-00614584.

XX PA (PROT-) PROTEIN DESIGN LABS INC.

XX PI Schneider WP, Vazquez M, Queen CL;

XX PT WPI; 2000-292433/25.

XX PR New fusion protein useful for treatment of autoimmune disorders comprises  
 XX a functional moiety of the extracellular domain of a Fas ligand protein  
 XX and a polypeptide capable of specifically binding to a cell surface  
 XX marker.

XX PS Disclosure; Col 35-36; 26pp; English.

CC The invention relates to a fusion protein comprising the Fas protein  
 CC recognition moiety of the extracellular domain of a Fas ligand (FasL)  
 CC protein and a protein capable of specifically binding to a cell surface  
 CC marker. The cell surface-binding protein may, for example, be the  
 CC variable region of an antibody, or it could comprise a ligand for a  
 CC cellular receptor (e.g., a growth factor or a cytokine), the ligand  
 CC binding domain of a cellular receptor, or the binding site of a cellular  
 CC adhesion molecule (e.g., an integrin). The fusion protein additionally  
 CC has an amino acid substitution in the FasL protein that decreases the  
 CC ability of the FasL portion of the fusion protein to form dimers,  
 CC trimers, or other aggregates. FasL is believed to act as a trimer,  
 CC transmitting an apoptotic signal via binding to its receptor, Fas.  
 CC Binding of a FasL trimer to two or more Fas molecules causes Fas to  
 CC dimerise, which transmits an apoptotic signal to the Fas-expressing cell.  
 CC The FasL portion of the fusion protein of the invention is less able to  
 CC aggregate and thereby cause apoptosis of Fas-expressing cells than free  
 CC soluble FasL. However, it is able to cause apoptosis of those cells which  
 CC are also able to bind the cell-surface binding portion of the fusion  
 CC protein. The fusion protein is therefore specifically cytotoxic for



XX	Pain-relieving target construct, linker moiety #12.
DE	
XX	Pain relief; tumour; therapeutic; breast cancer; lung cancer;
KW	Hodgkin's lymphoma; neuroblastoma; ischaemic.
XX	
OS	Synthetic.
XX	
FH	Key Location/Qualifiers
FT	Misc-difference 3..57 /note= "Each tetra-glycine repeat can optionally be di-
FT	glycine or tri-glycine"
FT	Misc-difference 8..57 /note= "Optionally absent"
FT	Misc-difference 58 /note= "Any amino acid"
PN	MO200147512-A2.
PD	05-JUL-2001.
XX	
PF	06-DEC-2000; 2000WO-US042661.
XX	
PR	08-DEC-1999; 99US-00457498.
XX	
PA	(FLUO-) FLUORO PROBE INC.
XX	
PI	Lutken GA;
DR	WPI; 2001-465198/50.
PT	Treatment of pain associated with an interior disease site, involves administering a pain-relieving target construct to the patient.
XX	
PS	Disclosure; Page 21; 31pp; English.
XX	
CC	The sequence represents the amino acid sequence of linker moiety #12 used to make a pain-relieving target construct. The construct comprises a pain-relieving agent linked to a ligand moiety that selectively binds to or is taken up by the tissue associated with the painful interior disease site. The construct is used for treating pain associated with an interior disease site. Since the pain-relieving agent is delivered by the ligand to the disease site intractable pain situated in the interior of the body such as caused by various tumours, such as breast cancer, lung cancer, Hodgkin's and non-Hodgkin's lymphoma, and neuroblastoma, and ischaemic CC and diseased tissues can be managed using a lower level of the pain relieving agent than is required when the pain-relieving agent is injected in the free state
CC	
XX	
SO	Sequence 60 AA;
Query Match	100.0%; Score 196; DB 4; Length 60;
Best Local Similarity	100.0%; Pred. No. 1e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 35; Conservative	0; Mismatches 0; Indels 0; Gaps 0;
OY	1 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 35       Db 3 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 37
RESULT 7	
ID	AAG77867
AC	AAG77867 standard; protein; 60 AA.
XX	
AA	AAG77867;
XX	
DT	08-MAY-2002 (first entry)
XX	
DE	MHC class II H domain spacer.
XX	
KM	MHC class II H domain spacer; IG3g-HLA fusion protein;
KW	major histocompatibility complex; MHC-peptide complex; MHC;
KM	MHC class I alpha chain; beta-2 microglobulin; MHC class II alpha chain;

KM	MHC class II beta chain; vaccine; immune response modulation;
KM	hyperproliferative disorder; neoplasm; hypergammaglobulinaemia;
KM	viral infection; hepatitis; meningitis; bacterial infection;
KM	tuberculosis; gingivitis; parasitic infection; autoimmune disease;
KM	Hashimoto's disease; Graves' disease; rheumatoid arthritis; allergy;
KM	asthma; organ rejection; graft-versus-host disease; GVHD; HLA;
KM	human leukocyte antigen.
XX	
OS	Synthetic.
XX	
PN	WO200178768-A2.
XX	
PD	25-OCT-2001.
XX	
PF	12-APR-2001; 2001WO-US011912.
XX	
PR	12-APR-2000; 2000US-0196472P.
XX	
PA	(UVRP ) UNIV ROCHESTER.
XX	
PI	Zauderer M, Smith ES;
XX	
DR	WPI, 2001-602927/68.
XX	
PT	Novel compound comprising major histocompatibility complex-peptide
PT	complexes, used to modulate immune responses.
XX	
PS	Example 3; Page 146; 16pp; English.
XX	
CC	The present sequence represents a major histocompatibility complex (MHC)
CC	class II H domain spacer. This protein was used as a replacement for the
CC	H domain present in MHC class II molecules in an example of the invention
CC	determining the construction of human IgG3-HA fusion proteins. Absence of
CC	the H domain or a suitable spacer would result in misfolding of the MHC
CC	class II molecule and the correct MHC binding site would not be formed.
CC	The invention comprises a compound which contains one or more major
CC	histocompatibility complex (MHC)-peptide complexes, and an antibody
CC	specific for a cell surface marker. The complexes comprise an MHC class I
CC	alpha chain, a beta-2 microglobulin molecule and an antigenic peptide
CC	bound in the MHC groove. Alternatively, the complexes may comprise an MHC
CC	class II alpha chain, an MHC class II beta chain, and an antigenic
CC	peptide bound in the MHC groove. The complexes are linked to the carboxyl
CC	terminus of the antibody. The compounds of the invention can be used as a
CC	vaccine to modulate an immune response. The compounds of the invention
CC	are useful for treating hyperproliferative disorders (e.g. neoplasms and
CC	hypergammaglobulinaemia); viral infections (e.g. hepatitis and
CC	meningitis); bacterial infections (e.g. tuberculosis and gingivitis);
CC	parasitic infections; autoimmune diseases (e.g. Hashimoto's disease,
CC	Graves' disease and rheumatoid arthritis); allergic reactions/conditions
CC	(e.g. asthma). The compounds of the invention may also be used in the
CC	treatment of organ rejection or graft-versus-host disease (GVHD)
XX	
SEQ	Sequence 60 AA:
Query Match	100.0%; Score 196; DB 4; Length 60;
Best Local Similarity	100.0%; Pred. No. 1e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
Qy	1 GCGSGCGSGCGSGCGSGCGSGCGSGCGSGCGSGCGS 35
Dd	1 GCGSGCGSGCGSGCGSGCGSGCGSGCGSGCGSGCGS 35
RESULT 8	
ID AAM18564	
ID AAM18564 standard; protein; 312 AA.	
AC AAM18564;	
XX	
DT 27-MAR-1998 (first entry)	
XX	
DE Novel fusion protein KDTF7.	
XX	

KW	Kunitz type active site inhibitor domain; fusion protein; Factor VIIa;
KV	FVIIa; linker; tissue factor; TF; treatment; procoagulant activity;
KM	thrombosis; human.
XX	
OS	Homo sapiens.
XX	
XX	Synthetic.
PH	
FT	Key Location/Qualifiers
FT	Domain 1..58
FT	/note= "Kunitz type active site inhibitor domain"
FT	Domain 59..93
FT	/note= "linker peptide"
FT	Domain 94..312
FT	/note= "tissue factor domain"
XX	
PN	WO9720939-A1.
PD	
PD	12-JUN-1997.
XX	
PE	22-NOV-1996; 96WO-USO18756.
PR	01-DEC-1995; 95US-00566459.
PR	04-DEC-1995; 95US-00566800.
XX	
PA	(GETH ) GENENTECH INC.
XX	
E1	Kelley RF, Lazarus RA, Lee GF;
DR	WPI, 1997-319785/29.
PT	
PT	Factor VIIa active site inhibitor domain and tissue factor fusion protein
PT	- used for inhibiting human tissue factor-FVIIa procoagulant activity in
PS	mammals, to treat e.g. thrombosis and restenosis.
XX	
PS	Example 2; Page; 58pp; English.
XX	
CC	This is a novel fusion protein comprising a factor VIIa (FVIIa) Kunitz
CC	type active site inhibitor domain sequence (AAW18553)', a linker domain
CC	sequence, and a tissue factor (TF) domain sequence which is a human
CC	tissue factor variant hFPAA (AAW18556). This novel fusion protein has a
CC	Kunitz domain linked at its C-terminus via the peptide linker to the N-
CC	terminus of hPTA. The novel fusion protein can be used for inhibiting
CC	human tissue factor-FVIIa (TF-FVIIa) procoagulant activity in a mammal.
CC	It can be used to treat TF-FVIIa related diseases like chronic
CC	thromboembolic diseases or disorders associated with fibrin formation
CC	including vascular disorders such as deep venous thrombosis, arterial
CC	thrombosis, stroke, tumour metastasis, thrombolysis, arteriosclerosis and
CC	restenosis following angioplasty. The novel fusion protein can also be
CC	used to treat acute and chronic disorders such as inflammation, septic
CC	shock, adult respiratory distress syndrome, septicemia, hypotension,
CC	disseminated intravascular coagulopathy and other diseases. Fusion
CC	proteins in which a FVIIa active site inhibitor and a tissue factor are
CC	covalently linked are more potent inhibitors of FVIIa than the individual
CC	components alone or a non-covalent mixture of the two. Note: This
CC	sequence is not explicitly shown in the specification; it has been
CC	created by combining a Kunitz domain sequence TP71-C (AAW18553)', a linker
CC	sequence given in Page 43 followed by a hTPAA sequence (AAW18556)
XX	
SQ	Sequence 312 AA;
XX	
Query Match	100.0%; Score 196; DB 2; Length 312;
Best Local Similarity	100.0%; Pred. No. 3.9e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
Gy	1 GCGGCGGCGGCGGCGGCGGCGGCGGCGGCGS 35
Dd	59 GGGGSGGGGCGGCGGCGGCGGCGGCGGCGGCGS 93
RESULT 9	
AAB00158	
ID	AAB00158 standard; protein; 507 AA.
XX	

AC	AA800158;
XX	
DT	08-FEB-2001 (first entry)
XX	
DE	sCD4-SCFv(17b) HIV single chain antibody fusion protein.
XX	
KM	Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;
KW	acquired immune deficiency syndrome; neutralisation; infection;
KW	gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;
KW	binding domain; single chain antibody; chimera; chimeric protein.
XX	
OS	Human immunodeficiency virus.
XX	Synthetic.
PN	WO200055207-A1.
PD	21-SEP-2000.
PE	16-MAR-2000; 2000WO-US006946.
PR	16-MAR-1999; 99US-0124681P.
PA	(USSH ) US NAT INST OF HEALTH.
PI	Berger EA, Del Castillo CM;
DR	WP1; 2000-638183/61.
DR	N-PSTDB; AAA54045.
PT	
PT	Novel neutralizing bispecific fusion proteins effective in viral such as
PT	HIV neutralization, comprises two different binding domains, inducing-
PS	binding domain and induced-binding domain functionally linked by linker.
XX	
PS	Claim 39; Page 46-47; 55pp; English.
CC	
CC	sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of
CC	binding to two sites of its target protein. The protein comprises a first
CC	binding domain capable of binding to an inducing site on the target
CC	protein, a second binding domain capable of forming neutralising complex
CC	with an induced epitope of the target protein and a linker connecting the
CC	binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment
CC	(containing domains D1 and D2) fused to a single chain Fv portion of
CC	antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or
CC	mimetic is used for inactivating gp120 protein of HIV, and for
CC	neutralising HIV. It is also used for blocking and preventing the binding
CC	of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte
CC	CD4 and for inhibiting HIV replication. The chimeric proteins is
CC	therefore useful for treating HIV infection and also AIDS. It is are
CC	particularly useful in the prevention of infection during or immediately
CC	after HIV exposure (e.g., mother/infant transmission, post-exposure
CC	prophylaxis, and as a topical inhibitor) and for providing long term
CC	resistance to HIV infections and AIDS. Gene therapy is used to secrete
CC	the bispecific protein at mucosal surfaces, such as the vaginal, rectal
CC	or oral mucosa. The fusion proteins is highly potent, broadly cross-
CC	reactive with neutralising antibody with high in vivo activity and no Pc-
CC	mediated undesirable targeting properties. When the fusion protein is
CC	substantially derived from human proteins, it has minimal immunogenicity
CC	and toxicity in humans which is of great value in prevention of infection
CC	during or immediately after HIV exposure
XX	
SO	Sequence 507 AA;
Query Match	100.0%; Score 196; DB 3; Length 507;
Best Local Similarity	100.0%; Pred. No. 5.7e-11;
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0	
OY	1 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 35       Db 209 GGCGSGGCGSGGCGSGGCGSGGCGSGGCGSGGCGS 243

ID AAB00159 standard; peptide; 39 AA.  
 XX AAB00159;  
 XX  
 DT 08-FEB-2001 (first entry)  
 DE Seven repeat linker.  
 KM Fusion protein; HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
 KM acquired immune deficiency syndrome; neutralisation; infection;  
 KM gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
 KM binding domain; single chain antibody; chimera; chimeric protein.  
 XX  
 OS Synthetic.  
 XX  
 PN W0200055207-A1.  
 PD 21-SEP-2000.  
 XX 16-MAR-2000; 2000MO-US006946.  
 PF 16-MAR-1999; 99US-0124681P.  
 PR (USSH ) US NAT INST OF HEALTH.  
 PA Berger EA, Del Caestillo CM;  
 PI WPI; 2000-638183/61.  
 XX N-PSDB; AAA54046.  
 DR  
 PT Novel neutralizing bispecific fusion proteins effective in viral such as  
 PT HIV neutralization, comprises two different binding domains, inducing-  
 PT binding domain and induced-binding domain functionally linked by linker.  
 XX  
 PS Example 1; Page 48; 55pp; English.  
 XX  
 CC sCD4-SCFv(17b) is a neutralising bispecific fusion protein capable of  
 CC binding to two sites of its target protein. The protein comprises a first  
 CC binding domain capable of binding to an inducing site on the target  
 CC protein, a second binding domain capable of forming neutralising complex  
 CC with an induced epitope of the target protein and a linker connecting the  
 CC binding domains. sCD4-SCFv(17b) comprises a soluble CD4 fragment  
 CC (containing domains D1 and D2) fused to a single chain Fv portion of  
 CC antibody 17b via a linker. sCD4-SCFv(17b), its variant, analogue or  
 CC mimetic is used for inactivating gp120 protein of HIV, and for  
 CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure. To generate the fusion protein  
 CC the full length CD4 coding sequence was excised from the plasmid pCB-3  
 CC using restriction enzymes StuI and SpeI. Synthetic oligonucleotides  
 CC (AAA54049, AAA54046) were ligated and used to regenerate the cleaved StuI  
 CC site and the next two bases of the CD4 CDNA and to produce an SpeI  
 CC overhang at the 3' end. The sequence also reconstructs the second domain  
 CC of CD4 and encodes the 37 amino acid intermediate linker  
 CC (Gly 4Ser)<sub>6</sub>Gly 4Thr 2Ser followed directly by the universal translation  
 CC termination signal. A BamHI site is included near the end of the third  
 CC (Gly 4Ser)<sub>6</sub> repeat to enable linkage to the 17b SCFv with the exact linker  
 CC sequence. The resulting intermediate plasmid was designated pCD1 but the  
 CC sCD4 linker insert was then subcloned into a pSCS3 derivative lacking a  
 CC BamHI site, forming pCD2. The 17b SCFv sequence was amplified from its  
 CC starting plasmid vector pmt del 0 using two primers (AAA54047, AAA54048).

CC The amplified fragment was digested with BamHI and SpeI and force-cloned  
 CC into pCD2 to generate the sCD4-SCFv(17b) fusion protein in a plasmid  
 CC designated pCD3  
 XX  
 SQ Sequence 39 AA;  
 Query Match 98.5%; Score 193; DB 3; Length 39;  
 Best Local Similarity 97.1%; Pred. No. 1,4e-11;  
 Matches 34; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 35  
 DB 3 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGT 37  
 RESULT 11  
 ID AB009677 standard; peptide; 40 AA.  
 XX  
 AC AB009677;  
 XX  
 DT 02-JUL-2003 (first entry)  
 DE INF-beta-fragment of crystallisation fusion protein related linker #6.  
 XX  
 KM Interferon-beta; INF-beta; fragment of crystallisation; Fc; IgG4;  
 KM cytosarctic; virucide; hepatocytic; antiinflammatory; tumour;  
 KM viral infection; interferon; INF; immunoglobulin Fc; hepatitis B;  
 KM lymphoma; leukaemia; linker.  
 XX  
 OS Synthetic.  
 XX  
 PN US2003026779-A1.  
 PD 06-FEB-2003.  
 XX 03-DEC-2001; 2001US-00005438.  
 PF 15-OCT-1999; 99US-00418734.  
 PR (YUL/) YU L.  
 PA (CHAN) CHANG T W.  
 PI Yu L, Chang TW;  
 XX WPI; 2003-417234/39.  
 DR  
 XX Treating tumors or viral infections, involves administering a hybrid  
 PT molecule comprising an interferon molecule joined at one end to a first  
 PT end of an immunoglobulin Fc fragment, with or without a linker.  
 XX  
 PS Example 2; Page 3; 9pp; English.  
 XX  
 CC The invention describes a method of treating tumors or viral infections,  
 CC involving administering a hybrid molecule having an interferon (IFN)  
 CC molecule (I) or its variants, joined at one end to a first end of an  
 CC immunoglobulin Fc fragment (II), without any linker between (I) and (II)  
 CC or with a first linker between (I) and (II). The method is useful for  
 CC treating tumors and viral infections including hepatitis B. The IFN-Fc  
 CC hybrids have a much more longer half-life in vivo than the native  
 CC IFNalpha. They are suitable for treating tumors such as lymphomas and  
 CC leukemias, as these products have a long retention time in the  
 CC vasculature and will not penetrate undesired sites. The hybrids have been  
 CC shown to ablate tumors in an animal model. At equivalent molar dosage of  
 CC native IFN and the hybrid, the hybrid provides a several hundred fold  
 CC efficacy at the same dosage and less frequent administration. This is the  
 CC amino acid sequence of a linker that can be used to create an interferon-  
 CC beta (INF-beta)-IgG4 fragment of crystallisation (Fc) fusion protein  
 XX  
 SQ Sequence 40 AA;  
 Query Match 89.8%; Score 176; DB 6; Length 40;  
 Best Local Similarity 88.6%; Pred. No. 5.1e-10;



XX	antithyroid; antidiabetic; neuroprotective; nociceptic; osteopathic;
KW	antirheumatic; antirheumatic; vaccine; immunisation; infectious disease;
KW	anti-viral protection; tumour; allergy; drug addiction; Crohn's disease;
KW	grief-verus-host disease; Grave's disease; diabetes; multiple sclerosis;
KW	Alzheimer's disease; osteoporosis; rheumatoid arthritis;
XX	inflammatory autoimmune disease.
XX	
XX	Synthetic.
XX	
XX	WO2003024480-A2.
XX	
XX	27-MAR-2003.
XX	
XX	16-SEP-2002; 2002MO-IB004252.
XX	
XX	14-SEP-2001; 2001US-0318967P.
XX	
XX	(CYTO-) CYTOS BIOTECHNOLOGY AG.
XX	
XX	
XX	Bachmann MF, Storni T, Lechner F;
XX	
XX	WPI; 2003-363095/34.
XX	
XX	A composition, useful for enhancing an immune response against an antigen
XX	or a virus-like particle, enhancing anti-viral protection in an animal,
XX	or immunizing or treating tumors or infectious diseases, e.g. viral
XX	infections.
XX	
XX	Disclosure; Page 64; 243pp; English.

CC The present invention describes a composition (C) for enhancing an immune  
CC response against an antigen or a virus-like particle in an animal. (C)  
CC comprises a virus-like particle (VLP) bound to at least one antigen, or a  
CC VLP capable of being recognised by the immune system of the animal. Also  
CC described: (1) enhancing an immune response against an antigen or a VLP  
CC in an animal comprising introducing (C) into the animal; (2) vaccines  
CC comprising (C) together with a pharmaceutical diluent, carrier or  
CC excipient; (3) immunising or treating an animal comprising administering  
CC the vaccine to the animal, or priming or boosting a T cell response in  
CC the animal by administering the vaccine; and (4) enhancing anti-viral  
CC protection in an animal comprising introducing (C) into the animal. (C)  
CC has cytostatic, virucide, antibacterial, antiparasitic, fungicide,  
CC antiallergic, immunosuppressive, antidiabetic, antiinflammatory,  
CC antithyroid, antidiabetic, neuroprotective, nootropic, osteopathic,  
CC antirheumatic and antiarthritic activities. (C) or the vaccines can be  
CC used for enhancing an immune response against an antigen or a VLP in an  
CC animal, enhancing anti-viral protection in an animal, or immunising or  
CC treating tumours and infectious diseases such as viral, bacterial,  
CC parasitic or fungal infections. The vaccine compositions are also useful  
CC for preventing or treating allergies, drug addiction, graft-versus-host  
CC disease, Crohn's disease, Grave's disease, diabetes, multiple sclerosis,  
CC Alzheimer's disease, osteoporosis, rheumatoid arthritis, or inflammatory  
CC autoimmune disease. ACC659838 to ACC659652 and ABR564901 to ABR565509  
CC represent sequences used in the exemplification of the present invention  
XX Sequence 41 AA;

[illegible]

RESULT	15
ABR44505	
ID	ABR44505 standard; peptide; 41 AA
AC	
XX	ABR44505;
XX	
DT	25-JUL-2003 (first entry)

XX	Peptide linker #4.
DE	
XX	Immunostimulatory; virus-like particle; bacteriophage; HBV; LCMV;
KW	hepatitis B virus; lymphocytic choriomeningitis virus; vaccine;
KW	hepatitis B virus; cytostatic; antiallergic; virucide; antibacterial;
KW	immunostimulant; cytostatic; antiallergic; virucide; antibacterial;
KW	immune response; immunisation; allergy; tumour; breast cancer;
KW	neuroblastoma; leukaemia; viral disease; influenza; hepatitis; measles;
KW	chicken pox; bacterial infection; tuberculosis; pneumonia; syphilis.
XX	
OS	Synthetic.
XX	
PN	WO2003024481-A2.
XX	
PD	27-MAR-2003.
XX	
PE	16-SEP-2002; 2002WO-IB004132.
XX	
PR	14-SEP-2001; 2001US-0318994P.
XX	22-APR-2002; 2002US-0374145P.
XX	
PA	(CYTO-) CYTOS BIOTECHNOLOGY AG.
PA	(MAUR-) MAURER P.
PA	(TISS-) TISSOT A.
PA	(SCHW-) SCHWARZ K.
PA	(MEIJ-) MEIJERINK E.
PA	(LIPO-) LIPOWSKY G.
PA	(PUMP-) PUMPENS P.
PA	(CIEL-) CIELENS I.
PA	(RENH-) RENHOFFA R.
XX	
PI	Maurer P, Tissot A, Schwarz K, Meijerink E, Lipowsky G;
PI	Pumpens P, Cie lens I, Renhofs R, Bachmann WF, Storni T;
XX	
DR	WPI; 2003-354564/33.
XX	
PT	New compositions comprising immunostimulatory substances packaged into
PT	virus-like particles, useful as a vaccine for enhancing an immune
PT	response in animals, e.g. for treating or preventing allergies, tumors or
PT	viral infections.
XX	
PS	Disclosure; Page 75; 322p; English.

The present invention describes a composition (C) for enhancing an immune response in an animal. (C) comprises a virus-like particle (VLP), and an immunostimulatory substance. The immunostimulatory substance is bound to the VLP. Also described: (1) enhancing an immune response in an animal by introducing (C) into the animal; (2) producing (C) for enhancing an immune response in an animal; (3) vaccines comprising (C) together with a pharmaceutical diluent, carrier or excipient; and (4) immunising or treating an animal by: (a) administering the vaccine to the animal; (b) priming a T cell response in the animal by administering the vaccine; or (c) boosting a T cell response in the animal by administering the vaccine. (C) has immunostimulant, cytostatic, antiallergic, vitucide and antibacterial activities. (1) can be used in vaccines for enhancing an immune response in an animal, particularly a mammal or human. Specifically, (C) is useful for enhancing a B cell response, a T cell response, or a cytotoxic T-lymphocyte (CTL) response. (C) or a vaccine comprising (C) can also be used for immunising or treating an animal, e.g. humans, sheep, horses, cattle, pigs, dogs, cats, rats, birds, reptiles or fish. (C) is particularly useful in prophylactic or therapeutic vaccines against allergies, tumours (e.g. breast cancers, neuroblastoma, or leukaemia), viral diseases (e.g. influenza, hepatitis, measles or chicken pox), or bacterial infections (e.g. tuberculosis, pneumonia or syphilis). ACC69790 to ACC69815 and ABR44502 to ABR44612 represent sequences used in the exemplification of the present invention

Query Match	87.8%	Score 172;	DB 6;	Length 41;
Best Local Similarity	90.9%	Pred. No. 1.2e-09;		
Matches 30; Conservative	0;	Mismatches 3;	Indels 0;	Gaps 0;



**THIS PAGE BLANK (USPTO)**

OTHER INFORMATION: present or absent?

Sequence 4, April  
Sequence 4, April  
Sequence 4, April  
Sequence 82, April  
Sequence 8319, April  
Sequence 90, April  
Sequence 8322, April  
Sequence 11, April  
Sequence 50, April  
Sequence 50, April  
Sequence 50, April  
Sequence 11, April  
Sequence 11, April  
Sequence 50, April  
Sequence 50, April  
Sequence 9, April  
Sequence 50, April  
Sequence 194, April





RESULT 9  
 US-09-054-281-22  
 Sequence 22, Application US/09054281  
 Patent No. 644421  
 GENERAL INFORMATION:  
 APPLICANT: Chung, Jay H.  
 TITLE OF INVENTION: Methods for Detecting Intermolecular  
 TITLE OF INVENTION: Interactions In Vivo and In Vitro  
 NUMBER OF SEQUENCE ADDRESSES: 22  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: Townsend and Townsend and Crew LLP  
 STREET: Two Embarcadero Center, Eighth Floor  
 CITY: San Francisco  
 STATE: California  
 COUNTRY: USA  
 ZIP: 94111-3834  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: PatentIn Release #1.0, Version #1.30  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/054,281  
 FILING DATE: 02-APR-1998  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 60/080,234  
 FILING DATE: 03-APR-1997  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/826,622

```

RESULT 10
US-09-478-948-6
Sequence 6, Application US/09478948
Patent No. 6465258
GENERAL INFORMATION:
APPLICANT: Shan, Bel
APPLICANT: Okamoto, Arthur Y.
APPLICANT: Tularik Inc.
TITLE OF INVENTION: FXR Receptor-Mediated Modulation of Cholesterol
FILE REFERENCE: 018781-001310US
CURRENT APPLICATION NUMBER: US/09/478,948
CURRENT FILING DATE: 2000-01-06
PRIOR APPLICATION NUMBER: US 60/115,249
PRIOR FILING DATE: 1999-01-07
NUMBER OF SEQ ID NOS: 6
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 6
LENGTH: 201
TYPE: PRT
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence:flexible linker
NAME/KEY: MOD_RES
LOCATION: (1)..(97)
OTHER INFORMATION: Gly at positions 1-97 may be present or absent
NAME/KEY: MOD_RES
LOCATION: (105)..(201)
OTHER INFORMATION: Gly at positions 105-201 may be present or absent
US-09-478-948-6

```

Query Match	85.7%	Score 168;	DB 4;	Length 201;
Best Local Similarity	82.4%;	Pred. No. 6.4e-09;		
Matches 28; Conservative	0;	Mismatches 6;	Indels 0;	Gaps 0;



TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-835-099A-9

Query Match 77.6%; Score 152; DB 2; Length 268;  
Best Local Similarity 63.6%; Pred. No. 2.2e-07;  
Matches 28; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

Qy 1 GGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGG 34  
Db 12 GGGGGGGGGGIGGGIGNVLGGLISGAGGGGGGGGGGGGG 55

## RESULT 14

US-09-157-349-9  
Sequence 9, Application US/09157349  
Patent No. 6068990  
GENERAL INFORMATION:  
APPLICANT: SHINTANI, Yasuaki  
APPLICANT: NISHI, Kazuo  
APPLICANT: KAWAMOTO, Tomohiro  
TITLE OF INVENTION: NOVEL PROTEINS, THEIR PRODUCTION  
TITLE OF INVENTION: AND USE  
NUMBER OF SEQUENCES: 18  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSMAN, LLP  
STREET: 130 Water Street  
City: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/157,349  
FILING DATE:  
CLASSIFICATION:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/835,099  
FILING DATE:  
APPLICATION NUMBER: 97105508.2  
FILING DATE: 03-APR-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Reenick, David S  
REGISTRATION NUMBER: 34,235  
REFERENCE/DOCKET NUMBER: 47342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
TELEX:  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 268 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-09-157-349-9

Query Match 77.6%; Score 152; DB 3; Length 268;  
Best Local Similarity 63.6%; Pred. No. 2.2e-07;  
Matches 28; Conservative 1; Mismatches 5; Indels 10; Gaps 1;

Qy 1 GGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGG 34  
Db 12 GGGGGGGGGGIGGGIGNVLGGLISGAGGGGGGGGGGGGG 55

RESULT 15  
US-09-046-985-15  
Sequence 15, Application US/09046985  
Patent No. 6121236  
GENERAL INFORMATION:  
APPLICANT: Ben-Sasson, Shmuel A.  
TITLE OF INVENTION: NOVEL MULTIVALENT LIGANDS WHICH MODULATE  
TITLE OF INVENTION: ANGIOGENESIS  
NUMBER OF SEQUENCES: 15  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: Hamilton, Brook, Smith & Reynolds, P.C.  
STREET: Two Militia Drive  
City: Lexington  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02173  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/046,985  
FILING DATE:  
CLASSIFICATION:  
ATTORNEY/AGENT INFORMATION:  
NAME: Brook, David E.  
REGISTRATION NUMBER: 22,592  
REFERENCE/DOCKET NUMBER: CMCC-614  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (781) 861-6240  
TELEFAX: (781) 861-9540  
INFORMATION FOR SEQ ID NO: 15:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 190 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-046-985-15

Query Match 76.5%; Score 150; DB 3; Length 190;  
Best Local Similarity 40.2%; Pred. No. 2.5e-07;  
Matches 35; Conservative 0; Mismatches 0; Indels 52; Gaps 2;

Qy 1 GGGGGGGGGGGG-----GGGGGGGGGGGGGGGGGG 30  
Db 27 GGGGGGGGGGGGSDFTAYRWRLSHRPKDLYSIVRRADRGGGGSKDFT 86

Qy 31 -----GGGG 35  
Db 87 AYRWRLSHRPKDLYSIVRRADRGGGG 113  
Search completed: November 21, 2004, 13:44:44  
Job time : 5.53467 sec

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Comugen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:42:32 ; Search time 15.3285 Seconds  
(without alignments)  
808.591 Million cell updates/sec

Title: US-09-936-702-2

Perfect score: 196  
Sequence: 1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSG 35

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1570615 seqs, 354127592 residues

Total number of hits satisfying chosen parameters: 1570615

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:\*

- 1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/2/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/2/pubpaa/US10\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/2/pubpaa/US10D\_PUBCOMB.pep.\*
- 17: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/2/pubpaa/US11\_NEW\_PUB.pep.\*
- 19: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*
- 20: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	196	100.0	35	US-10-013-173-30	Sequence 30, Appl
2	196	100.0	35	US-10-150-752-30	Sequence 30, Appl
3	196	100.0	35	US-10-244-821-30	Sequence 30, Appl
4	196	100.0	40	US-10-081-400-1	Sequence 1, Appl
5	196	100.0	40	US-10-257-864A-137	Sequence 137, App
6	196	100.0	40	US-10-257-864A-138	Sequence 138, App
7	196	100.0	40	US-10-608-710-6	Sequence 6, Appl
8	196	100.0	40	US-10-768-873-1	Sequence 1, Appl
9	196	100.0	50	US-09-948-039-32	Sequence 32, Appl
10	196	100.0	60	US-09-832-297A-12	Sequence 12, Appl
11	196	100.0	60	US-09-833-203-18	Sequence 18, Appl
12	182	92.9	211	US-10-437-963-126553	Sequence 12653,
13	179.5	91.6	250	US-10-425-115-288960	Sequence 288960,

14	176	89.8	40	14	US-10-005-438-9	Sequence 9, Appl
15	173.5	88.5	133	17	US-10-425-115-250884	Sequence 250884,
16	173.5	88.5	156	15	US-10-634-862-118955	Sequence 11, Appl
17	173	88.3	184	17	US-10-425-115-288955	Sequence 288955,
18	172	87.8	394	15	US-10-433-108-26	Sequence 26, Appl
19	172	87.8	302	15	US-10-433-108-25	Sequence 25, Appl
20	171	87.2	160	17	US-10-739-930-7816	Sequence 7816, Ap
21	168	85.7	30	10	US-09-833-203-19	Sequence 19, Appl
22	168	85.7	200	9	US-09-798-584-18	Sequence 18, Appl
23	168	85.7	200	9	US-09-967-624-19	Sequence 19, Appl
24	168	85.7	200	9	US-09-998-667-18	Sequence 18, Appl
25	168	85.7	200	10	US-09-921-159-14	Sequence 34, Appl
26	168	85.7	200	10	US-09-990-940-21	Sequence 21, Appl
27	168	85.7	200	10	US-09-989-981A-13	Sequence 13, Appl
28	168	85.7	200	10	US-09-850-948-29	Sequence 29, Appl
29	168	85.7	200	14	US-10-160-354-4	Sequence 4, Appl
30	168	85.7	200	14	US-10-026-021-8	Sequence 8, Appl
31	168	85.7	200	14	US-10-161-165-3	Sequence 3, Appl
32	168	85.7	200	14	US-10-160-663-3	Sequence 3, Appl
33	168	85.7	200	14	US-10-071-838-15	Sequence 15, Appl
34	168	85.7	200	14	US-10-094-417-25	Sequence 25, Appl
35	168	85.7	200	14	US-10-188-405-13	Sequence 13, Appl
36	168	85.7	200	14	US-10-273-575-29	Sequence 29, Appl
37	168	85.7	200	14	US-10-233-098-5	Sequence 5, Appl
38	168	85.7	200	14	US-10-245-850-3	Sequence 3, Appl
39	168	85.7	200	14	US-10-293-582-27	Sequence 27, Appl
40	168	85.7	200	14	US-10-100-818-14	Sequence 14, Appl
41	168	85.7	200	14	US-10-237-467-18	Sequence 18, Appl
42	168	85.7	200	14	US-10-321-204-54	Sequence 54, Appl
43	168	85.7	200	14	US-10-179-766-12	Sequence 12, Appl
44	168	85.7	200	14	US-10-123-568-4	Sequence 4, Appl
45	168	85.7	200	14	US-10-123-731-10	Sequence 10, Appl

## ALIGNMENTS

RESULT 1  
US-10-013-173-30  
Sequence 30, Application US/10013173  
Publication No. US20030095977A1  
GENERAL INFORMATION:  
APPLICANT: Goshorn, Stephen C.  
APPLICANT: Graves, Scott Stoll  
APPLICANT: Schultz, Joanne Elaine  
APPLICANT: Lin, Yukang  
APPLICANT: Sanderson, James A.  
APPLICANT: Reno, John M.  
TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
TITLE OF INVENTION: METHODS OF USE THEREOF  
FILE REFERENCE: 690022.547C1  
CURRENT APPLICATION NUMBER: US/10/013.173  
CURRENT FILING DATE: 2001-12-07  
NUMBER OF SEQ ID NOS: 69  
SOFTWARE: FastSeq for Windows Version 4.0  
SEQ ID NO 30  
LENGTH: 35  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Linker sequence  
US-10-013-173-30

Query Match 100.0%; Score 196; DB 14; Length 35;  
Best Local Similarity 100.0%; Pred. No. 2.6e-09;  
Matches 35; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSG 35  
DB 1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSG 35

RESULT 2

RESULT 4  
US-10-081-400-1

```

1 RESULT 5
2 US-10-257-864A-137
3 ; Sequence 137, Application US/10257864A
4 ; Publication No. US20040058393A1
5 ; GENERAL INFORMATION:
6 ; APPLICANT: FUKUSHIMA, NAOSHI
7 ; APPLICANT: TSUCHIYA, MASAYUKI
8 ; APPLICANT: OH-EDA, MASAYOSHI
9 ; APPLICANT: UNO, SHINSUKE
10 ; APPLICANT: KIKUCHI, YASUFUMI
11 ; APPLICANT: GOTEIBA-SHI, OTOMO
12 ; TITLE OF INVENTION: AGONIST ANTIBODIES
13 ; FILE REFERENCE: 065678-0107
14 ; CURRENT APPLICATION NUMBER: US/10/257,864A
15 ; CURRENT FILING DATE: 2003-07-24
16 ; PRIOR APPLICATION NUMBER: PCT/JP01/03288
17 ; PRIOR FILING DATE: 2001-04-17
18 ; PRIOR APPLICATION NUMBER: PCT/JP01/01912
19 ; PRIOR FILING DATE: 2001-03-12
20 ; PRIOR APPLICATION NUMBER: JP2000-115246
21 ; PRIOR FILING DATE: 2000-04-17
22 ; PRIOR APPLICATION NUMBER: JP2000-321821
23 ; PRIOR FILING DATE: 2000-10-20
24 ; PRIOR APPLICATION NUMBER: JP2000-321822
25 ; PRIOR FILING DATE: 2000-10-20
26 ; NUMBER OF SEQ ID NOS: 138
27 ; SOFTWARE: PatentIn Ver. 2.1
28 ; SEQ ID NO 137
29 ; LENGTH: 40
30 ; TYPE: PRT
31 ; ORGANISM: Artificial Sequence
32 ; FEATURE:
33 ; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
34 ; OTHER INFORMATION: peptide
35 ; FEATURE:
36 ; NAME/KEY: PEPTIDE
37 ; LOCATION: (1)..(40)
38 ; OTHER INFORMATION: This peptide may range from 1-40 amino acids:
39 ; OTHER INFORMATION: Refer to the specification as filed for preferred
40 ; OTHER INFORMATION: emodiments

```



```

; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Gly-Ser linker
; OTHER INFORMATION: that may encompass 5 to 50 residues, preferably
; OTHER INFORMATION: 5 to 20 residues.
; OS-09-949-039-32

```

Query Match	100.0%	Score 196;	DB 10	Length 50;
Best Local Similarity	100.0%	Pred. No. 3.3e-09;		
Matches 35;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

**Qy**

**Db**

**1**

**1**

**3.5**

**3.5**

```

RESULT 10
US-09-832-297A-12
Sequence 12, Application US/09832297A
Publication No. US20010055566A1
GENERAL INFORMATION:
APPLICANT: Fluoroprobe, Inc.
APPLICANT: LUIKEN, George A.
TITLE OF INVENTION: METHOD FOR VIEWING TUMOR TISSUE LOCATED WITHIN A BODY CAVITY
FILE REFERENCE: FLUOR1120-2
CURRENT APPLICATION NUMBER: US/09/832,297A
PRIORITY FILING DATE: 2001-04-09
PRIORITY FILING DATE: 1999-07-28
PRIORITY FILING DATE: 1999-07-28
PRIORITY FILING DATE: 1999-10-15
PRIORITY FILING DATE: 1999-10-15
NUMBER OF SEQ ID NOS: 12
SOFTWARE: PatentIn version 3.1
SEQ ID NO 12
LENGTH: 60
TYPE: PRF
ORGANISM: Artificial sequence
FEATURE:
OTHER INFORMATION: Peptide linker moiety
FEATURE:
NAME/KEY: REPEAT
LOCATION: (3)..(6)
OTHER INFORMATION: Amino Acid at residue 3 could be repeated up to 3 times
FEATURE:
NAME/KEY: REPEAT
LOCATION: (3)..(57)
OTHER INFORMATION: Amino Acids at residues 3 to 7 could be repeated up to 10 times
US-09-832-297A-12

```

Query Match	100.0%	Score 196 ; DB 9 ;	Length 60 ;
Best Local Similarity	100.0%	Pred. No. 3.8e-09 ;	
Matches 35 ; Conservative	0 ;	Mismatches 0 ;	Indels 0 ; Gaps 0 ;

Qy 1 35  
Db 3 37

RESULT 11  
US-09-833-203-18  
Sequence 18 Application US/098333203  
Publication No. US20030166277A1  
GENERAL INFORMATION:  
APPLICANT: Zauderer, Maurice  
APPLICANT: Smith, Ernest S.  
TITLE OF INVENTION: Targeted Vaccine Delivery System  
FILE REFERENCE: 1821.0020001  
CURRENT APPLICATION NUMBER: US/09/833,203  
CURRENT FILING DATE: 2001-04-12  
PRIOR APPLICATION NUMBER: US 60/196,472  
PRIOR FILING DATE: 2000-04-12  
NUMBER OF SEQ ID NOS: 63

```

; SOFTWARE: PatentIn version 3.0.0
; SEQ ID NO 18
; LENGTH: 60
; TYPE: BRT
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; OTHER INFORMATION: Spacer
US-03-833-203-18

```

Query Match	100.0%	Score 196;	DB 10;	Length 60;
Best Local Similarity	100.0%	Pred. No. 3.8e-09;		
Matches 35;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;

[illegible]

```

RESULT 12
US-10-437-963-126553
Sequence 126553, Application US/10437963
Publication NO. US20040123343A1
GENERAL INFORMATION:
APPLICANT: La Rosa, Thomas J.
APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
APPLICANT: Wu, Wei
APPLICANT: Boukharov, Andrey A.
APPLICANT: Barbazuk, Brad
APPLICANT: Li, Ping
TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53221)B
CURRENT APPLICATION NUMBER: US/10/437, 963
CURRENT FILING DATE: 2003-05-14
NUMBER OF SEQ ID NOS: 204966
SEQ ID NO 126553
LENGTH: 211
TYPE: PRT
ORGANISM: Oryza sativa
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT4530_29086C.1.pep
US-10-437-963-126553

```

Query Match	92.9%	Score 182;	DB 16;	Length 211;
Best Local Similarity	94.1%	Pred. No. 1.1e-07;		
Matches 32;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0.

[illegible]

```

RESULT 13
US-10-425-115-288960
: Sequence 288960, Application US/10425115
: Publication No. US20040214272A1
: GENERAL INFORMATION:
: APPLICANT: La Rosa, Thomas J.
: APPLICANT: Kovalic, David K.
: APPLICANT: Zhou, Yihua
: APPLICANT: Cao, Yongwei
: TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
: TITLE OF INVENTION: Plants
: FILE REFERENCE: 38-21(53222)B
: CURRENT APPLICATION NUMBER: US/10/425,115
: CURRENT FILING DATE: 2003-04-28
: NUMBER OF SEQ ID NOS: 369326
: SEQ ID NO 288960
: LENGTH: 250
: TYPE: prt
/

```

```
Search completed: November 21, 2004, 14:03:51
Job time : 18.3285 secs
```

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:29:16 ; Search time 3.83212 Seconds  
(without alignments)  
878.780 Million cell updates/sec

Title: US-09-936-702-2

Perfect score: 196  
Sequence: 1 GGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGS 35

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 79:\*

1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	162.5	82.9	207	2 T07381	glycine-rich prote
2	162	82.7	221	2 T04592	glycine-rich cell
3	159.5	81.4	165	1 KNRZG1	glycine-rich cell
4	159	81.1	136	2 T29282	hypothetical prote
5	155	79.1	592	2 E82759	endo-1,4-beta-gluc
6	154.5	78.8	1226	2 T24045	hypothetical prote
7	154	78.6	266	1 CTRBL	calpain (EC 3.4.22
8	153	78.1	643	1 KRRU2	keratin 1, type II
9	152	77.6	268	1 CTRBL	calpain (EC 3.4.22
10	150.5	76.8	255	2 B84777	hypothetical prote
11	150	76.5	106	2 F84797	hypothetical prote
12	149.5	76.3	440	2 S71795	transcription fact
13	149	76.0	302	2 C84470	hypothetical prote
14	148.5	75.8	396	2 T49109	glycine-rich prote
15	148	75.5	259	2 T15126	hypothetical prote
16	147	75.0	481	2 A35628	loricrin - mouse
17	146.5	74.7	393	2 T20268	hypothetical prote
18	146	74.5	266	1 CTRBL	calpain (EC 3.4.22
19	145.5	74.2	271	2 S34666	glycine-rich prote
20	145	74.0	203	1 JQ1061	glycine-rich prote
21	145	74.0	385	2 T20410	hypothetical prote
22	145	74.0	404	2 S54729	RNA-binding protei
23	145	74.0	569	1 KRWSE1	keratin, 59k type
24	144.5	73.7	291	1 S31415	glycine-rich prote
25	144.5	73.7	581	1 KRWSE2	keratin, type II c
26	143.5	73.2	384	1 A26099	glycine-rich cell
27	143	73.0	139	2 S14443	glycine-rich RNA-b
28	143	73.0	158	2 T05254	probable RNA-bind
29	143	73.0	171	2 H84709	probable glycine-r

30	143	73.0	420	2 I59234	octamer binding tr
31	142	72.4	408	2 S57483	glycine-rich protei
32	142	72.4	420	2 A49642	transcription fact
33	142	72.4	561	2 A31994	keratin 10, type I
34	141.5	72.2	681	2 AB2155	hypothetical prote
35	141	71.9	183	2 PN0109	keratin-like prote
36	141	71.9	839	2 B96576	hypothetical prote
37	140.5	71.7	316	1 A38743	loricrin - human
38	140	71.4	570	2 S07330	keratin, epidermal
39	140	71.4	979	2 A35913	regulatory factor
40	139	70.9	252	1 S01821	glycine-rich prote
41	139	70.9	284	2 S74256	homeotic protein s
42	139	70.9	330	2 S74255	homeotic protein s
43	138.5	70.7	212	2 A44994	eggshell protein 1
44	138.5	70.7	561	2 T22917	probable ATP-depen
45	138	70.4	424	1 Z3BPF1	coat protein A pre

## ALIGNMENTS

## RESULT 1

T07381  
glycine-rich protein Tfms - tomato  
C.Species: Lycopersicon esculentum (tomato)  
C.Date: 14-May-1999 #sequence\_revision 14-May-1999 #text\_change 09-Jul-2004  
C.Accession: T07381  
R.Santino, C.G.; Stanford, G.L.; Conner, T.W.  
Plant Mol. Biol. 33, 405-416, 1997  
A.Title: Developmental and transgenic analysis of two tomato fruit enhanced genes.  
A.Reference number: Z16000; MUID:97201476; PMID:9049262  
A.Accession: T07381  
A.Status: preliminary; translated from GB/EMBL/DBJ  
A.Molecule type: DNA  
A.Residues: 1-207 <SAS>  
A.Cross-references: UNIPROT:Q43522; EMBL:X95262; NID:g1166449; PIDN:CAA64559.1; PID:g116  
A.Experimental source: cultivar UG22b; fruit  
C.Genetics:  
A.Gene: Tfms  
C.Superfamily: hydroxyproline-rich glycoprotein

Query Match 82.9%; Score 162.5; DB 2; Length 207;  
Best Local Similarity 81.1%; Pred. No. 4.6e-07;  
Matches 30; Conservative 0; Mismatches 4; Indels 3; Gaps 1;

QY 1 GGGSGGGGS---GGGSGGGSGGGSGGGSGGGSGGGG 34  
DB 84 GGGSGGGSGTGGGGSGGGSGGGSGGGSGGGGGGGG 120

## RESULT 2

T04592

glycine-rich cell wall structural protein homolog F23E13.120 - Arabidopsis thaliana  
C.Species: Arabidopsis thaliana (mouse-ear cress)  
C.Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 09-Jul-2004  
C.Accession: T04592  
R.Bevan, M.; Halbert, H.; Braun, M.; Holzer, E.; Brandt, A.; Duesterhoeft, A.; Jeeze, T.  
Submitted to the Protein Sequence Database, March 1998  
A.Reference number: Z15378  
A.Accession: T04592  
A.Molecule type: DNA  
A.Residues: 1-221 <BEV>  
A.Cross-references: UNIPROT:O65514; EMBL:AL022141  
A.Experimental source: cultivar Columbia; BAC clone F23E13  
C.Genetics:  
A.Map position: 4  
A.Note: F23E13.120

Query Match 82.7%; Score 162; DB 2; Length 221;  
Best Local Similarity 69.8%; Pred. No. 5.3e-07;  
Matches 30; Conservative 0; Mismatches 5; Indels 8; Gaps 1;  
QY 1 GGGSGGGSGGGSGGGSGGGSGGGSGGGSGGGSGGS 35

[illegible]

### RESULT 3

KNRZG1  
 glycine-rich cell wall structural protein 1 precursor (clone lambda-313) - rice  
 C:Species: *Oryza sativa* (rice)  
 C:Date: 31-Dec-1991 #sequence\_revision 31-Dec-1991 #text\_change 09-Jul-2004  
 C:Accession: S13385  
 R:Lei, M.; Wu, R.  
 Plant Mol. Biol. 16, 187-198, 1991  
 A:Title: A novel glycine-rich cell wall protein gene in rice.  
 A:Reference number: S13385; MUID:91370862; PMID:1716496  
 A:Accession: S13385  
 A:Molecule type: DNA  
 A:Releases: 1-165 <LEI>  
 A:Cross-references: UNIPROT:P25074; EMBL:X53596; NID:G20246; PIDN:CAA37665.1; PID:G20247  
 C:Genetics:  
 A:Gene: grp-1  
 C:Superfamily: glycine-rich cell wall structural protein 1  
 C:Keywords: cell wall; duplication; structural protein  
 F:1-23/Domain: signal sequence #status predicted <Sig>  
 F:24-165/Product: glycine-rich cell wall structural protein 1 #status predicted <Mat>  
 F:30-55/Region: repeat R1  
 F:56-62/Region: repeat R2  
 F:62-92/Region: repeat R1  
 F:93-99/Region: repeat R2  
 F:100-131/Region: repeat R1  
 F:132-138/Region: repeat R2  
 F:139-160/Region: repeat R1

```
Query Match      81.4%   Score 159.5   DB 1; length 165;  
Best Local Similarity    62.5%   Pred No. 6.6e-07;  
Matches        30; Conservative    0; Mismatches    5; Indels    13; Gaps    1.  
  
Oy       1 GGGSGGCGSGGCGSG-----GGGSGGCGSGGCGSGGCGGS 35  
          |||||  
db       111 GGGGGGGGGGGGGGGGCGSGGCGGYGAKCGGGGGGGGGCGGGGG 158  
          |||||
```

## RESULT 4

T29282  
hypothetical protein C34D4.11 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T29282  
R:Du, Z.; Le, T.T.  
submitted to the EMBL Data Library, May 1996  
A:Description: The sequence of C. elegans cosmid C34D4.  
A:Reference number: Z20600  
A:Accession: T29282  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-136 <DUZ>  
A:Cross-references: UNIPROT:Q18444; EMBL:U58755; PINN:AA00696.1; GSPDB:GN000222; CESP:CT  
A:Experimental source: strain Bristol N2; clone C34D4

	Query March	81.1%	Score 159;	DB 2;	Length 126;
	Best Local Similarity	79.4%	Pred. No. 6.2e-07;		
	Matches 27;	Conservative 1;	Mismatches 6;	Indels 0;	Gaps 0.
Oy	1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGG	34			
Ob	85 GGGGNGGGGRGGGGGGGGGGGGGGGGGGGGG	118			

## RESULT 5

```
E82759
endo-1,4-beta-glucanase XF0818 [imported] - Xylella fastidiosa (strain 9a5c)
C:Species: Xylella fastidiosa
C:Date: 18-Aug-2000 #sequence_revision 20-Aug-2000 #next_change 09-Jul-2004
```

C/Accession: E82759  
R/Anonymous: The Xylella fastidiosa Consortium for Nucleotide Sequencing  
Nature 406, 151-157, 2000  
A/Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A/Reference number: A82515; MUID:20365717; PMID:10910347  
A/Note: for a complete list of authors see reference number A59328 below  
A/Accession: E82759  
A/Status: preliminary  
A/Molecule type: DNA  
A/Residues: 1-592 <STM>  
A/Cross-references: UNIPROT:Q9PFE0; GB:AE003921; GB:AE003849; NID:G9105710; PIDN:AAF8362E  
A/Experimental source: strain 945c  
R/Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Briotes, W.R.S.; Bueno, M.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carret, H  
de-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferreira, A.J.S.  
A/Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Frohme  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, U.P.; Krieger, J.E.; Kurame, E.E.; Laigret  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.U.; Marques, M.V.; Martins, B  
A/Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miraccea, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasaki  
A/Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveira  
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Ze  
A/Reference number: A59328  
A/Contents: annotation  
C/Genetic8:  
/Gene: XF0818

	Query Match	79.1%	Score 155;	DB 2:	Length 592;
	Beat Local Similarity	74.4%;	Pred. No.	4.1e-06;	
	Matches 23;	Conservative 1;	Mismatches 5;	Indels 4;	Gaps 1;
Qy	1 GGGGSGGGSSGGCGSGGGSSGCAG--SACCGSGGGGS	35			
b					
bb	467 GGGGSGGGSSGGCGSGGGSSGCAGSCTGCCGGS	505			

## RESULT

T24045  
hypochlorite protein R08B4.1 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 09-Jul-2004  
C:Accession: T24045  
R:White, S.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: Z19834  
A:Accession: T24045  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1226 <WIL>  
A:Cross-References: UNIPROT:Q21335, EMBL:Z68008, PIDD:CAA92000.1, GSPDB:GN00028, CESP:ROT  
A:Experimental source: clone R08B4

```
Query Match Similarity 78.8%; Score 154.5; DB 2; Length 1226;
Best Local Similarity 71.8%; Pred. No. 88-06;
Matches 28; Conservative 3; Mismatches 3; Indels 5; Gaps 1;
```

OY            2 CCGSGGGGSCGGGSCGGGS-----CCGGSGGGGSCGGGCS 35  
             | | | | | | | | | | : | | | | | | | |  
Db            864 GGGGGGGGGGSCGGSGGGGSSNSNGGGGGGCGGGGCGGCN 902

## RESULT

CIRBL

calpain (EC 3.4.22.17) small chain - rabbit  
N:Alternate names: calcium-activated neutral proteinase (CANP); calpain light chain; cal  
C:Species: *Oryctolagus cuniculus* (domestic rabbit)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
C:Accession: A24816  
J:Emori, Y.; Kawasaki, H.; Imajo, S.; Kawashima, S.; Suzuki, K.  
J. Biol. Chem. 261, 9472-9476, 1986  
A:Title: Isolation and sequence analysis of cDNA clones for the small subunit of rabbit  
A:Reference number: A24816; MUID:86250903; PMID:3013892  
A:Accession: A24816  
A:Molecule type: mRNA  
A:Residues: 1-266 <E0>  
A:Cross-references: UNIPROT:P06813; GB:M13364; NID:G164875; PIDN:AAA81565.1; PID:G164876  
C:Complex: heterodimer of L (large) and S (small) chains  
C:Function:  
A:Description: catalyzes the hydrolysis of peptides  
A>Note: cleaves preferentially after tyrosine, methionine, or arginine residues and befo  
C:Superfamily: calpain small chain; calmodulin repeat homology  
C:Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyd  
F:1-54/Domain: glycine-rich <GLY>  
F:94-125/Domain: calmodulin repeat homology <EF1>  
F:137-169/Domain: calmodulin repeat homology <EF2>  
F:170-199/Domain: calmodulin repeat homology <EF3>  
F:202-234/Domain: calmodulin repeat homology <EF4>  
F:235-266/Domain: calmodulin repeat homology <EF5>

Query Match 78.6%; Score 154; DB 1; Length 266;  
Best Local Similarity 62.2%; Pred. No. 2.6e-06;  
Matches 28; Conservative 3; Mismatches 4; Indels 10; Gaps 1;

1 GGGGSGGGSGGGG-----GGSGGGSGGGSGGGSGGGG 35  
Db 11 GGGGSGGGGGLGGGLGNVLTGLISGAGGGGGGGGGGAGGGGT 55

RESULT 8  
KRU02  
keratin 1, type II, cytoskeletal - human  
N:Alternate names: 67K type II epidermal keratin; cytokeratin 1  
C:Species: *Homo sapiens* (man)  
C:Date: 04-Dec-1986 #sequence\_revision 22-Oct-1999 #text\_change 10-Dec-1999  
C:Accession: A22940; A02950; A43342  
R:Johnson, L.D.; Idler, W.W.; Zhou, X.M.; Roop, D.R.; Steinert, P.M.  
Proc. Natl. Acad. Sci. U.S.A. 82, 1896-1900, 1985  
A:Reference number: A22940; MUID:85166339; PMID:2580302  
A:Accession: A22940  
A:Molecule type: DNA  
A:Residues: 1-643 <J0H>  
A:Cross-references: GB:M98776; GB:M1215; GB:M11845; GB:M11846; NID:G1843461  
A>Note: translation of initiator Met 18 not shown  
R:Steinert, P.M.; Parry, D.A.D.; Idler, W.W.; Johnson, L.D.; Steven, A.C.; Roop, D.R.  
J. Biol. Chem. 260, 7142-7149, 1985  
A:Title: Amino acid sequences of mouse and human epidermal type II keratins of M-r 67,00  
late filament subunits.  
A:Reference number: A92535; MUID:85207740; PMID:2581964  
A:Accession: A02950  
A:Molecule type: mRNA  
A:Residues: 151-183, 'K', 185-199, 'M', 201-204, 'K', 206-236, 'S', 238-239, 'R', 241-356, 'Y', 358-  
'S', 638-643 <STE>  
A:Cross-references: GB:M10938; NID:G186787; PIDN:AAA36153.1; PID:G386854  
A:Experimental source: tissue neonatal foreskin  
A>Note: the authors translated the codon CUG for residue 476 as Met  
R:Chipev, C.C.; Kotze, B.P.; Markova, N.; Bale, S.J.; Digiovanna, J.J.; Compton, J.G.; S  
Cell 70, 821-828, 1992  
A:Title: A leucine----proline mutation in the H1 subdomain of keratin 1 causes epidermol  
A:Reference number: A43342; MUID:92286601; PMID:1381288  
A:Accession: A43342  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 144-146, 'P', 148-159, 'P', 161-183, 'K', 185-186 <CHI>  
A:Cross-references: GB:M98776; GB:M1215; GB:M11845; GB:M11846; NID:G1843461  
A>Note: sequence extracted from NCBI backbone (NCBI:112784)  
C:Comment: The cytoskeletal and microfibrillar keratins are classified into two types, B

atin IF protein subunit appears to be a heterotrimer of two type I and two type II pro  
C:Comment: Keratin 1 is expressed in terminally differentiating epidermis.  
C:Genetics:  
A:Gene: GDB:KRT1  
A:Cross-references: GDB:128198; OMIM:139350  
A:Map position: 12q11-12q13  
A>Note: defects in this gene may result in epidermolytic hyperkeratosis  
C:Complex: heterotrimer of two type I, usually keratin 10 (see PIR:KRU0), and two typ  
C:Superfamily: cytoskeletal keratin  
C:Keywords: coiled coil; heterotrimer; intermediate filament  
F:4-179/Domain: head <HED>  
F:14-143/Region: E1 and V1 subdomains  
F:14-179/Region: H1 subdomain  
F:180-492/Domain: rod <ROD>  
F:180-214/Region: coil 1A  
F:215-226/Region: linker 1A  
F:227-327/Region: coil 1B  
F:328-344/Region: linker 12  
F:345-363/Region: coil 2A  
F:364-371/Region: linker 2  
F:372-492/Region: coil 2B  
F:430/Region: stutler  
F:493-643/Domain: tail <END>  
F:493-512/Region: H2 subdomain  
F:513-643/Region: V2 and E2 subdomains

Query Match 78.1%; Score 153; DB 1; Length 643;  
Best Local Similarity 82.4%; Pred. No. 6.3e-06;  
Matches 28; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

1 GGGGSGGGSGGGSGGGSGGGSGGGSGGGG 34  
Db 93 GGGGFGGGFGGGGFGGGGFGGGGFGGGGFGGGG 126

RESULT 9  
CIH01  
calpain (EC 3.4.22.17) small chain - human  
N:Alternate names: calcium-activated neutral proteinase (CANP)  
C:Species: *Homo sapiens* (man)  
C:Date: 28-Dec-1987 #sequence\_revision 28-Dec-1987 #text\_change 09-Jul-2004  
C:Accession: A26107; A23650  
R:Miyaake, S.; Emori, Y.; Suzuki, K.  
Nucleic Acids Res. 14, 8805-8817, 1986  
A:Title: Gene organization of the small subunit of human calcium-activated neutral prote  
A:Reference number: A93648; MUID:87066759; PMID:3024120  
A:Accession: A26107  
A:Molecule type: DNA  
A:Residues: 1-268 <MY>  
A:Cross-references: UNIPROT:P04632; GB:M31502  
R:Ohno, S.; Emori, Y.; Suzuki, K.  
Nucleic Acids Res. 14, 5559, 1986  
A:Title: Nucleotide sequence of a cDNA coding for the small subunit of human calcium-dep  
A:Reference number: A93631; MUID:86286563; PMID:3016651  
A:Accession: A23650  
A:Molecule type: mRNA  
A:Residues: 1-268 <OH>  
A:Cross-references: EMBL:X04106; NID:G35327; PIDN:CAA27726.1; PID:G35328  
C:Genetics:  
A:Gene: GDB:CANP4  
A:Cross-references: GDB:119752; OMIM:114170  
A:Map position: 19pter-19pter  
A:Introns: 70/2; 81/3; 111/3; 131/1; 152/3; 175/3; 202/1; 241/1; 260/3  
C:Complex: heterodimer of L (large) and S (small) chains  
C:Function:  
A:Description: catalyzes the hydrolysis of peptides  
A>Note: cleaves preferentially after tyrosine, methionine, or arginine residues and befo  
C:Superfamily: calpain small chain; calmodulin repeat homology  
C:Keywords: calcium binding; cysteine proteinase; duplication; EF hand; heterodimer; hyd  
F:1-56/Domain: glycine-rich <GLY>  
F:96-127/Domain: calmodulin repeat homology <EF1>  
F:133-171/Domain: calmodulin repeat homology <EF2>  
F:172-201/Domain: calmodulin repeat homology <EF3>



A:Experimental source: cultivar Columbia; BAC clone FIN20  
 C:Genetics:  
 A:Gene: ATSP:AT4g22020  
 A:Map position: 4  
 C:Superfamily: Phaseolus glycine-rich cell wall protein 1.8

Query Match 75.8%; Score 148.5; DB 2; Length 396;  
 Best Local Similarity 53.7%; Pred. No. 9,7e-06;  
 Matches 29; Conservative 0; Mismatches 6; Indels 19; Gaps 1;

QY 1 GGGSGGGSGG-----GGSGGGSGGGSGGGSGGGGS 35  
 |||||  
 Db 251 GGGGGGGGGGSKVGGYGHSGFGGAGVFNSGGGGGGGGGGGGGGGNGS 304

## RESULT 15

T15126  
 hypothetical protein T20B6.3 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 20-Sep-1999 #sequence\_revision 20-Sep-1999 #text\_change 09-Jul-2004

C:Accession: T15126

R:Beck, C.; Wamsley, P.

submitted to the EMBL Data Library, April 1997

A:Description: The sequence of C. elegans cosmid T20B6.

A:Reference number: Z18297

A:Accession: T15126

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-259 <BEC>

A:Cross-references: UNIPROT:002049; EMBL:AF000193; NID:G1946986; PID:G1946988; PIDN:AAB5

A:Experimental source: strain Bristol N2; clone T20B6

C:Genetics:

A:Gene: CESP:T20B6.3

A:Map position: 3

A:Introns: 9/2; 231/1

C:Superfamily: Phaseolus glycine-rich protein 1.0

Query Match 75.5%; Score 148; DB 2; Length 259;

Best Local Similarity 77.8%; Pred. No. 7.6e-06;

Matches 28; Conservative 0; Mismatches 6; Indels 2; Gaps 1;

QY 1 GGGSGGGSGG--GGSGGGSGGGSGGGSGGGSGGGG 34  
 |||||  
 Db 184 GGGGFGGGMGGGGGGGGGGGGGGGGGGGGGGGG 219

Search completed: November 21, 2004, 13:43:27  
 Job time : 4.83212 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:28:31 ; Search time 19.461 Seconds  
(without alignments)  
1037.189 Million cell updates/sec

[illegible]

Scoring table: BLOSUM62 ,  
Gapop 10.0 , Gapext 0.5

Searched: 1825181 seqs, 575374646 residues  
Total number of hits satisfying chosen parameters: 1825181

```
Minimum DB seq length: 0
Maximum DB seq length: 20000000000
```

Post-processing:	Minimum Match	0%
	Maximum Match	100%
	Listing first	45 summaries

```
1: uniprot_sprot:*
2: uniprot_trembl:
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	172.5	88.0	738	2	002402	002402 pinctada f
2	169	86.2	551	2	075HG8	075HG8 oryza sativ
3	169	86.2	551	2	AAR89849	AAR89849 oryza sativ
4	162.5	82.9	207	2	Q43522	Q43522 lycopersicc
5	162	82.7	221	2	Q65514	Q65514 arabidopsis
6	161	82.1	192	2	Q928P7	Q928P7 rhizobium r
7	160	81.6	233	2	Q92N07	Q92N07 rhizobium r
8	159.5	81.4	165	1	GRP1.ORYSA	P25074 oryza sativ
9	159	81.1	136	2	RI18444	RI18444 caenorhabdi
10	158	80.6	296	2	ORU50	ORU500 arabidopsis
11	157	80.1	698	2	080KX8	080KX8 ectocarpus
12	156	79.6	397	2	06S9V4	06S9V4 musca domes
13	156	79.6	397	2	AAR32812	AAR32812 musca dom
14	156	79.6	527	2	06S9V3	06S9V3 musca domes
15	155	79.1	592	2	AAR32813	AAR32813 musca dom
16	155	79.1	527	2	09PF60	09PF60 xylella fast
17	155	79.1	1615	2	Q7RQW4	Q7RQW4 plasmodium
18	154.5	78.8	1160	2	Q7YXK8	Q7YXK8 caenorhabdi
19	154	78.6	161	2	Q95UX3	Q95UX3 drosophila
20	154	78.6	162	2	Q95NU5	Q95NU5 drosophila
21	154	78.6	163	2	Q95NU6	Q95NU6 drosophila
22	154	78.6	163	2	Q95UX4	Q95UX4 drosophila
23	154	78.6	164	2	Q95NP2	Q95NP2 drosophila
24	154	78.6	165	2	Q9GP44	Q9GP44 drosophila
25	154	78.6	165	2	Q95NR6	Q95NR6 drosophila
26	154	78.6	165	2	Q95UX2	Q95UX2 drosophila
27	154	78.6	266	1	CS51.RABIT	P736813 oryza sativ
28	154	78.6	284	2	Q73EF9	Q73EF9 bacillus ce
29	154	78.6	284	2	AAS39446	AAS39446 bacillus
30	154	78.6	340	2	081E8	081E8 bacillus ce
31	154	78.6	697	2	Q9GRW7	Q9GRW7 drosophila

32	153	78.1	643	1	K0264	homo sapien
33	153	78.1	644	2	AH63597	
34	152	77.6	167	2	Q9LQ28	Ah63597 homo sapi
35	152	77.6	268	1	CS51	arabidopsis
36	152	77.6	2359	2	081519	h calpain s
37	151	77.0	337	2	Q9ET75	plasmodium
38	151	77.0	718	2	Q91TR1	rattus norv
39	150.5	76.8	255	2	Q9SIH2	tupaia heip
40	150.5	76.8	774	2	086884	arabidopsis
41	150.5	76.8	774	2	CAD88219	caenorhabd
42	150.5	76.8	774	2	CAD88221	caenorhab
43	150	76.5	106	2	Q7XJ7	arabidopsis
44	149.5	76.3	440	1	FXGA	gallus gall
45	149	76.0	267	2	Q6D12	xenopus tro

## ALIGNMENTS

RESULT 1	002402	PRELIMINARY;	PRT;	738 AA.
ID	002402;			
AC	002402;			
DT	01-JUL-1997 (TrEMBLrel. 04, Created)			
DT	01-JUL-1997 (TrEMBLrel. 04, last sequence update)			
DT	01-OCT-2003 (TrEMBLrel. 25, last annotation update)			
DE	Insoluble protein.			
OS	Pinctada fucata.			
OC	Eukaryota; Metazoa; Mollusca; Bivalvia; Pteriomorphia; Pterioidea;			
OC	Pterioidea; Pteridae; Pinctada.			
OX	NCBI_TaxID=50426;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RX	MEDLINE=97320490; Pubmed=9177341;			
RA	Sudo S., Fujikawa T., Nagakura T., Ohkubo T., Sakaguchi K., Tanaka M.,			
RA	Nakashima K., Takahashi T.,			
RT	"Structures of mollusc shell framework proteins."			
RL	Nature 387:563-564 (1997).			
EMBL	EMBL; D86074; BAA20466.1;			
SD	SEQUENCE 738 AA; 61723 MW; PDP984139BF3BA59 CRC64;			

Query Match:	88.0%;	Score 172.5;	DB 2;	Length 738;
Best Local Similarity	78.6%;	Pred. No. 1.3e-06;		
Matches 33;	Conservative	0;	Mismatches 2;	Indels 7;
				Gaps 1;

[illegible]

R	Accession #
Q75HGB	
ID	Q75HGB
AC	Q75HGB
DT	05-JUL-2004 (TREMBLrel. 27, Created)
DT	05-JUN-2004 (TREMBLrel. 27, Last sequence update)
DT	05-JUL-2004 (TREMBLrel. 27, Last annotation update)
DE	Pulative prohibitin.
GN	Name=OSUJBA0024Fp18.27;
OS	Oryza sativa (japonica cultivar-group).
OC	Eukaryotes; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta
OC	Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
CC	Ehharctoidaeae; Oryzeae; Oryza.
OX	NCBI_TaxID=39947;
RN	[1]
RP	SEQUENCE FROM N.A.
RA	Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.
RA	Overton II L.L., Tsatirlin T., Kim M.M., Bera J.J., Jin S.S.,
RA	Fedorosh D.W., Tallon L.J., Koo H., Zismann V., Hsieh J., Blunt S.
RA	Vanaken S.S., Riedmuller S.B., Uttechkar T.T., Feldlyum T.V.,
RA	Yang Q.Q., Haas B.J., Sun B.B., Peterson J.J., Quackenbush J.,
RL	White O., Salzberg S.L., Frazer C.M.;
	Submitted (OCT-2002) to the EMBL/GeneBank/DBSJ databases.

```

RN [2]
RP SEQUENCE FROM N.A.
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135594; AAR9849.1; -.
DR InterPro; IPR001107; Band_7.
DR SMART; SMO0244; PHB; 1.
SQ SEQUENCE 551 AA; 59192 MW; 34C1EC87BB3CA297 CRC64;

Query Match      86.2%; Score 169; DB 2; Length 551;
Best Local Similarity 82.9%; Pred. No. 1.9e-06;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 GCGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGG 35
    |||||
DB 47 GGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGG 81

RESULT 3
AAR9849 PRELIMINARY; PRT; 551 AA.
AC AAR9849;
DT 02-MAR-2004 (TREMBLrel. 27, Created)
DT 02-MAR-2004 (TREMBLrel. 27, Last sequence update)
DT 02-MAR-2004 (TREMBLrel. 27, Last annotation update)
DE Putative Prohibitin.
GN OSUNBA0024F18.27.
OC Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrharioideae; Oryzaceae; Oryza; Oryza sativa.
CX NCBI_TaxID=39947;
[1]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;
RA Buell C.R., Yuan Q., Ouyang S., Liu J., Gansberger K., Jones K.M.,
RA Overton II L.L., Teltrin T., Kim M.M., Bera J.J., Jin S.S.,
RA Padroen D.W., Tallon L.J., Koo H., Zismann V., Hsiao J., Blunt S.,
RA Vanaken S.S., Riedmiller S.B., Utehrack T.T., Feldblyum T.V.,
RA Yang Q.Q., Haas B.J., Suh B.B., Peterson J.J., Quackenbush J.,
RA White O., Salzberg S.L., Fraser C.M.;
RT "Oryza sativa chromosome 3 BAC OSUNBA0024F18 genomic sequence.";
RT Submitted (OCT-2002) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RP STRAIN=cv. Nipponbare;
RA Buell R.;
RL Submitted (JAN-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC135594; AAR9849.1; -.
SQ SEQUENCE 551 AA; 59192 MW; 34C1EC87BB3CA297 CRC64;

Query Match      86.2%; Score 169; DB 2; Length 551;
Best Local Similarity 82.9%; Pred. No. 1.9e-06;
Matches 29; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 1 GCGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGG 35
    |||||
DB 47 GGGGSGGGGSGGGGSGGGGSGGGGSGGGGSGGGG 81

RESULT 4
O43522 PRELIMINARY; PRT; 207 AA.
AC O43522;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE T1m5 Protein.
GN Name=T1m5;
OS Lycopersicon esculentum (Tomato).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids;
OC Lamiales; Solanales; Solanaceae; Solanum.

```

RX	NCHI_TaxID=4081;			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RQ	TISSUE=Fruit;			
RA	MEDLINE=97201476; PubMed=9049262;			
RT	Santino C.G., Stanford G.L., Conner T.W.;			
RL	"developmental and transgenic analysis of two tomato fruit enhanced genes.";			
RN	Plant Mol. Biol. 33:405-416(1997).			
RM	(2)			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Fruit;			
RA	Connor T.W.;			
RL	Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; X95262; CA64559.1; -.			
DR	PIR; T07381; T07381.			
DR	HSP; P24337; IHP.			
DR	InterPro; IPRO03612; AAI.			
DR	Pfam; PF00204; TRYF_alpha_amy1. 1.			
DR	SMART; SM00499; AAI_1.			
SQ	SEQUENCE 207 AA; 18265 MW; A50FE71FBEDAC4D9 CRC64;			
<hr/>				
Query Match	82.9%; Score 162.5; DB 2; Length 207;			
Best Local Similarity	81.1%; Pred. No. 2.8e-06;			
Matches 30; Conservative 0; Mismatches 4; Indels 3; Gaps 1.				
OY	1 GGCGSGGGGGS--GGCGSGGGGSGGGGSGGGGSGGC 34 			
Db	84 GGSGGCGGGGTGGGGGSGGGGSGGGGSGGGGSGGGG 120			
<hr/>				
RESULT 5				
ID O65514	PRELIMINARY;	PRT:	221 AA.	
AC O65514;				
DT 01-AUG-1998 (TREMBLrel. 07, Created)				
DT 01-AUG-1998 (TREMBLrel. 07, Last annotation update)				
DT 05-JUL-2004 (TREMBLrel. 27, Last annotation update)				
DE Putative glycine-rich cell wall protein.				
GN Name=F3JEI3.120; Synonyms=AT9436230;				
OS Arabidopsis thaliana (Mouse-ear cress);				
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;				
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;				
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.				
OX NCBI_TaxID=3702;				
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,				
RA Hohenseil J., Jesse T., Heijnen L., Vos P., Mewes H.W., Mayer K.,				
RA Schueller C.;				
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.				
RN [2]				
RP SEQUENCE FROM N.A.				
RA EU Arabidopsis sequencing project;				
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.				
RN [3]				
RP SEQUENCE FROM N.A.				
RA Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,				
RA Mewes H.W., Lemcke K., Mayer K.F.X.;				
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
RN [4]				
RP SEQUENCE FROM N.A.				
RA EU Arabidopsis sequencing project;				
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.				
DR EMBL; AL022141; CAA18129.1; -.				
DR EMBL; AL161589; CAB80294.1; -.				
DR PIR; T04592; T04592.				
DR InterPro; IPRO02952; Eggshe11.				
DR PRINTS; PR01228; EGSHDEL.				
SQ SEQUENCE 221 AA; 20344 MW; DE8DAE2E2C57A9F8 CRC64;				
<hr/>				
Query Match	82.7%; Score 162; DB 2; Length 221;			
Best Local Similarity	69.8%; Pred. No. 3.2e-06;			





Db 471 GGGGGGTGAGGGGGGGGGGGGGGGGGGGG 505

## RESULT 12

06S9V4 PRELIMINARY; PRT; 397 AA.  
AC 06S9V4;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXF.  
GN Name=dax;  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461853; AAR23812.1; -  
DR InterPro; IPR001275; DM\_DNA-binding.  
DR Pfam; PF00751; DM; 1.  
DR SMART; SM00301; DM; 1.  
DR PROSITE; PS40000; DM\_DOMAIN\_1; 1.  
DR PROSITE; PS50809; DM\_DOMAIN\_2; 1.  
SQ SEQUENCE 397 AA; 41701 MW; 38B49CED6931F8C4 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 397;  
Best Local Similarity 76.5%; Pred. No. 1.6e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

0y 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34  
Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 13

AAR23812 PRELIMINARY; PRT; 397 AA.  
AC AAR23812;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXF.  
GN Dsx.  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461853; AAR23812.1; -  
SQ SEQUENCE 397 AA; 41701 MW; 38B49CED6931F8C4 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 397;  
Best Local Similarity 76.5%; Pred. No. 1.6e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

0y 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34

Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 14

06S9V3 PRELIMINARY; PRT; 527 AA.  
AC 06S9V3;  
DT 05-JUL-2004 (TREMBlrel. 27, Created)  
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)  
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXM.  
GN Name=dax;  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461854; AAR23813.1; -  
DR InterPro; IPR001275; DM\_DNA-binding.  
DR Pfam; PF00751; DM; 1.  
DR SMART; SM00301; DM; 1.  
DR PROSITE; PS40000; DM\_DOMAIN\_1; 1.  
DR PROSITE; PS50809; DM\_DOMAIN\_2; 1.  
SQ SEQUENCE 527 AA; 55663 MW; A71E04E000A6D421 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 527;  
Best Local Similarity 76.5%; Pred. No. 2e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

0y 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34  
Db 192 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 225

## RESULT 15

AAR23813 PRELIMINARY; PRT; 527 AA.  
AC AAR23813;  
DT 02-MAR-2004 (TREMBlrel. 27, Created)  
DT 02-MAR-2004 (TREMBlrel. 27, Last sequence update)  
DT 02-MAR-2004 (TREMBlrel. 27, Last annotation update)  
DE DSXM.  
GN Dsx.  
OS Musca domestica (House fly).  
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Muscoidea;  
OC Muscidae; Musca.  
OX NCBI\_TaxID=7370;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX PubMed=14673649;  
RA Hediger M., Burghardt G., Siegenthaler C., Buser N.,  
RA Hilfinger-Kleiner D., Dubendorfer A., Bopp D.;  
RT "Sex determination in Drosophila melanogaster and Musca domestica  
converges at the level of the terminal regulator doublesex.";  
RL Dev. Genes Evol. 214:29-42(2004).  
DR EMBL; AY461854; AAR23813.1; -  
SQ SEQUENCE 527 AA; 55663 MW; A71E04E000A6D421 CRC64;

Query Match 79.6%; Score 156; DB 2; Length 527;  
Best Local Similarity 76.5%; Pred. No. 2e-05;  
Matches 26; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

0y 1 GGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 34

Sun Nov 28 09:38:24 2004

us-09-936-702-2.rup

Page 6

[illegible]

Search completed: November 21, 2004, 13:42:23  
Job time : 21.4161 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: November 21, 2004, 13:25:21 ; Search time 262.343 Seconds  
(without alignments)  
694.642 Million cell updates/sec

Title: US-09-936-702-3  
Perfect score: 2668  
Sequence: 1 MNRGVPRFHLVLQALLP.....GTRLEIKLVPRGSGHHHHH 508

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2002273 seqs, 358729299 residues

Total number of hits satisfying chosen parameters: 2002273

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq.23Sep04:\*  
1: geneseqp19808:\*  
2: geneseqp19908:\*  
3: geneseqp20008:\*  
4: geneseqp20018:\*  
5: geneseqp20028:\*  
6: geneseqp20038:\*  
7: geneseqp20038:\*  
8: geneseqp20048:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2651.5	99.4	507	3	AA800158
2	1064	39.9	295	2	AA807605
3	1064	39.9	318	2	AA807606
4	1064	39.9	394	2	AA807606
5	1064	39.9	394	3	AA807606
6	1064	39.9	416	3	AA807606
7	1064	39.9	436	3	AA807606
8	1064	39.9	458	3	AA807606
9	1064	39.9	458	4	AA807606
10	1064	39.9	458	4	AA807606
11	1064	39.9	458	7	AA807606
12	1064	39.9	458	7	AA807606
13	1064	39.9	458	8	AA807606
14	1064	39.9	473	7	AA807606
15	1064	39.9	474	3	AA807606
16	1064	39.9	481	3	AA807606
17	1064	39.9	481	3	AA807606
18	1064	39.9	481	3	AA807606
19	1064	39.9	519	2	AA807606
20	1064	39.9	616	3	AA807606
21	1064	39.9	631	1	AA807606
22	1064	39.9	631	3	AA807606
23	1064	39.9	631	3	AA807606
24	1064	39.9	631	3	AA807606
25	1064	39.9	631	3	AA807606

## ALIGNMENTS

26	1064	39.9	729	1	AA807606
27	1064	39.9	729	3	AA807606
28	1064	39.9	729	3	AA807606
29	1064	39.9	729	3	AA807606
30	1064	39.8	458	7	AA807606
31	1059	39.7	398	2	AA807606
32	1059	39.7	398	2	AA807606
33	1059	39.7	400	2	AA807606
34	1059	39.7	402	1	AA807606
35	1059	39.7	402	1	AA807606
36	1059	39.7	458	1	AA807606
37	1059	39.7	458	1	AA807606
38	1059	39.7	458	1	AA807606
39	1059	39.7	458	2	AA807606
40	1059	39.7	458	2	AA807606
41	1059	39.7	462	2	AA807606
42	1059	39.7	462	2	AA807606
43	1059	39.7	462	2	AA807606
44	1059	39.7	462	2	AA807606
45	1059	39.7	462	2	AA807606

RESULT 1  
AAB00158 standard; protein; 507 AA.  
ID AAB00158 standard; protein; 507 AA.  
AC AAB00158;  
XX  
XX  
DT 08-FEB-2001 (first entry)  
XX  
XX  
DE sCD4-SCFV(17b) HIV single chain antibody fusion protein.  
Fusion protein: HIV; human immunodeficiency virus; antibody; Fv; AIDS;  
acquired immune deficiency syndrome; neutralisation; infection;  
KW gene therapy; CD4; gp120; glycoprotein; resistance; vaccination;  
KW binding domain; single chain antibody; chimera; chimeric protein.  
OS Human immunodeficiency virus.  
OS Synthetic.  
PN WO200055207-A1.  
XX  
XX  
PD 21-SEP-2000.  
XX  
PF 16-MAR-2000; 2000MO-US006946.  
XX  
PR 16-MAR-1999; 99US-0124681P.  
XX  
PA (USSH ) US NAT INST OF HEALTH.  
XX  
PI Berger EA, Del Castillo CM;  
XX  
XX WPI; 2000-638183/61.  
DR N-PSDB; AA54045.  
XX  
PT Novel neutralizing bispecific fusion proteins effective in viral such as  
PT HIV neutralization, comprises two different binding domains, inducing-  
PT binding domain and induced-binding domain functionally linked by linker.  
XX  
XX Claim 39; Page 46-47; 55pp; English.  
XX  
XX sCD4-SCFV(17b) is a neutralising bispecific fusion protein capable of  
XX binding to two sites of its target protein. The protein comprises a first  
XX binding domain capable of binding to an inducing site on the target  
XX protein, a second binding domain capable of forming neutralising complex  
XX with an induced epitope of the target protein and a linker connecting the  
XX binding domains. sCD4-SCFV(17b) comprises a soluble CD4 fragment  
XX (containing domains D1 and D2) fused to a single chain Fv portion of  
XX antibody 17b via a linker. sCD4-SCFV(17b), its variant, analogue or  
XX mimetic is used for inactivating gp120 protein of HIV, and for

CC neutralising HIV. It is also used for blocking and preventing the binding  
 CC of the viral or recombinant gp120 protein to soluble CD4 or lymphocyte  
 CC CD4 and for inhibiting HIV replication. The chimeric proteins is  
 CC therefore useful for treating HIV infection and also AIDS. It is are  
 CC particularly useful in the prevention of infection during or immediately  
 CC after HIV exposure (e.g., mother/infant transmission, post-exposure  
 CC prophylaxis, and as a topical inhibitor) and for providing long term  
 CC resistance to HIV infections and AIDS. Gene therapy is used to secrete  
 CC the bispecific protein at mucosal surfaces, such as the vaginal, rectal  
 CC or oral mucosa. The fusion proteins is highly potent, broadly cross-  
 CC reactive with neutralising antibody with high in vivo activity and no Fc-  
 CC mediated undesirable targeting properties. When the fusion protein is  
 CC substantially derived from human proteins, it has minimal immunogenicity  
 CC and toxicity in humans which is of great value in prevention of infection  
 CC during or immediately after HIV exposure

XX Sequence 507 AA;

Query Match 99.4%; Score 2651.5; DB 3; Length 507;

Best Local Similarity 99.8%; Pred. No. 9e-128; Mismatches 0; Indels 1; Gaps 1;

Matches 507; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MNRGVPFRHLILVQLALPAAATQGGKVVIGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 DB 1 MNRGVPFRHLILVQLALPAAATQGGKVVIGKGDVVELTCTASQKKSIOFHMNSNOIK 60  
 QY 61 ILNGSGFLTKGSPKNDRAADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120  
 DB 61 ILNGSGFLTKGSPKNDRAADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120  
 QY 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCVTIIONOKKVEFKIDIVLAFOKASGGSGSGSGSGSGSGSGSGSGSGSGSGSGG 240  
 DB 181 TWTCVTIIONOKKVEFKIDIVLAFOKASGGSGSGSGSGSGSGSGSGSGSGSGSGSGG 240  
 QY 241 GGSQVOLLGSAEVKKRGSVSKVCSKSGDTFIRYSFTWNRQAQGGLEMMGRITITLDV 300  
 DB 241 GGSQVOLLGSAEVKKRGSVSKVCSKSGDTFIRYSFTWNRQAQGGLEMMGRITITLDV 300  
 QY 301 AHYAPHLOGRVITITADKSTSTVYLELNLSDDTAIVYFCAGVYGEADEGEYDNNKFLKH 360  
 DB 301 AHYAPHLOGRVITITADKSTSTVYLELNLSDDTAIVYFCAGVYGEADEGEYDNNKFLKH 360  
 QY 361 WGGQTLVTISGGSGSGSGSGSGSGSSELTQSPATLSVSPGERATLSCRASESVSDDLA 420  
 DB 361 WGGQTLVTISGGSGSGSGSGSGSGSSELTQSPATLSVSPGERATLSCRASESVSDDLA 420  
 QY 421 WYQCKPQQAARLLIYGASTRATGVPARFSGSGSAGFTLTISIQSDFAVYYCOQYNNW 480  
 DB 421 WYQCKPQQAARLLIYGASTRATGVPARFSGSGSAGFTLTISIQSDFAVYYCOQYNNW 480  
 QY 481 PPRYTFQGTLEIKLVPRSGHHNNH 508  
 DB 481 PPRYTFQGTLEIKLVPRSGHHNNH 508  
 QY 481 PPRYTFQGTLEIKLVPRSGHHNNH 508  
 DB 481 PPRYTFQGTLEIKLVPRSGHHNNH 508

RESULT 2  
 AAR07605  
 ID AAR07605 standard; protein; 295 AA.

XX AAR07605;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)

DE Plasmid T4/LFA-3/AD gene product with C-terminal encoding  
 DE phosphatidylinositol (PI) linkage signaling sequence.  
 XX  
 XX Plasma membrane binding affinity; micelle.

XX Unidentified.

OS Location/Qualifiers

XX Key

1..266 /label= T4/LFA-3/AD gene product

FT Protein

267..295

FT Peptide

/label= PI signal peptide

XX

MO9012099-A.

XX

18-OCT-1990.

PD

10-APR-1989;

XX

89US-00335688.

PR

10-APR-1989;

XX

89US-00335688.

XX

(BIOJ ) BIOGEN INC.

XX

Walner BP;

XX

WPI, 1990-334849/44.

DR

N-PSDB; AAQ06404.

XX

Phosphatidylinositol linkage signalling DNA sequence - derived from

PT lymphocyte function-associated antigen 3, used for prodn. of chimeric

PS proteins.

XX

Disclosure, Fig 4; 53pp; English.

XX

The signal sequence is attached downstream of the plasmid sequence

CC encoding a secreted protein, which will then produce proteins covalently

CC anchored to the cell surface in which they were produced. This can give

CC rise to plasma membrane binding, enhanced purifiability, micelle

CC formation etc. especially useful in the production of chimeric targeted

CC drugs, to produce micellar or liposomal delivery systems or in enhanced

CC purification and screening of cells, proteins or DNA libraries. (Updated

CC on 09-JAN-2003 to add missing OS field.) (updated on 25-MAR-2003 to

CC correct PA field.)

XX

Sequence 295 AA;

QY

Query Match 39.9%; Score 1064; DB 2; Length 295;

DB

Best Local Similarity 100.0%; Pred. No. 6.2e-47;

QY

Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

DB

1 MNRGVPFRHLILVQLALPAAATQGGKVVIGKGDVVELTCTASQKKSIOFHMNSNOIK 60

QY

61 ILNGSGFLTKGSPKNDRAADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120

DB

61 ILNGSGFLTKGSPKNDRAADSRSLMDQGNFPLIINKLIEDSDTYICEVEDQKEEVOL 120

QY

121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180

DB

121 LVFGLTANSDTHLQGSGLTLTLESPPGSSPSVQCSPPRGKNIQGGKTLVSQLELDQSG 180

QY

181 TWTCVTIIONOKKVEFKIDIVLAFOKAS 208

DB

181 TWTCVTIIONOKKVEFKIDIVLAFOKAS 208

RESULT 3

AAR07606

ID AAR07606 standard; protein; 318 AA.

XX AAR07606;  
 XX  
 XX 25-MAR-2003 (revised)  
 DT 09-JAN-2003 (revised)  
 DT 06-FEB-1991 (first entry)

DE Plasmid T4/LFA-3/2 gene product with C-terminal encoding  
DE phosphatidylinositol (PI) linkage signaling sequence.  
KW Plasma membrane binding affinity; micelle.  
XX Unidentified.  
OS  
XX  
XX  
FH Key Location/Qualifiers  
FT Protein 1..267  
FT /label= LFA-3 gene product  
FT 267..318  
FT Peptide /label= PI signal peptide  
XX  
XX  
XX WO9012099-A.  
XX  
XX 18-OCT-1990.  
XX  
XX 10-APR-1989; 89US-00335688.  
XX  
XX 10-APR-1989; 89US-00335688.  
XX  
XX (BIOJ ) BIOGEN INC.  
XX  
XX Wallner BP;  
XX  
XX WPI, 1990-334849/44.  
XX  
XX N-PSDB; AAQ06405.  
XX  
XX Phosphatidylinositol linkage signalling DNA sequence - derived from  
XX lymphocyte function-associated antigen 3, used for prodn. of chimeric  
XX proteins.  
XX  
XX Disclosure; Fig 5; 53pp; English.  
XX  
XX The signal sequence is attached downstream of the Plasmid sequence  
XX encoding a secreted protein, which will then produce proteins covalently  
XX anchored to the cell surface in which they were produced. This can give  
XX rise to plasma membrane binding, enhanced purifiability, micelle  
XX formation etc. especially useful in the production of chimeric targeted  
XX drugs, to produce micellar or liposomal delivery systems or in enhanced  
XX purification and screening of cells, proteins or DNA libraries. (Updated  
XX on 09-FAN-2003 to add missing OS field.) (Updated on 25-MAR-2003 to  
XX correct PA field.)  
XX  
SQ Sequence 318 AA;  
Query Match 39.9%; Score 1064; DB 2; Length 318;  
Best Local Similarity 100.0%; Pred. No. 6.6e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRFHLILVLOALPPATQGGKVVLGKKGDTVELTCTASOKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRFHLILVLOALPPATQGGKVVLGKKGDTVELTCTASOKSIOFHWKNSNOIK 60  
QY ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSDFHLLQGSGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFHLLQGSGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208

RESULT 4  
AA39825  
ID AA39825 standard; protein; 394 AA.  
XX  
XX  
XX AAY39825;  
XX

DT 03-DEC-1999 (first entry)  
XX  
XX Soluble human T4 protein.  
XX  
XX Soluble T4 protein; sT4; human; HIV; binding inhibitor; T4+ cell; AIDS;  
KW vaccine; immunisation; therapy.  
XX  
XX Homo sapiens.  
OS  
XX  
XX US5958678-A.  
XX  
XX 28-SEP-1999.  
XX  
XX 12-DEC-1994; 94US-00354452.  
XX  
XX 21-AUG-1986; 86US-00898587.  
XX  
XX 11-JUN-1991; 91US-00713564.  
XX  
XX 06-JUL-1992; 92US-00909021.  
XX  
XX (UYCO ) UNIV COLUMBIA NEW YORK.  
XX  
XX McDougal JS, Weiss R, Axel R, Littman DR, Maddon PJ, Chess L;  
XX  
XX WPI, 1999-561025/47.  
XX  
XX N-PSDB; AAZ20694.  
XX  
XX Human T4 protein inhibits HIV binding to T4 cells, useful for treating  
XX AIDS.  
XX  
XX Disclosure; Col 13-16; 58pp; English.  
XX  
XX This sequence represents the soluble human T4 protein of the invention.  
XX The soluble human T4 protein blocks the binding of HIV to T4+ cells and  
XX is therefore useful for the treatment of AIDS. Monoclonal antibodies  
XX against the T4 protein may be used as vaccines for immunising subjects  
XX against AIDS  
XX  
SQ Sequence 394 AA;  
Query Match 39.9%; Score 1064; DB 2; Length 394;  
Best Local Similarity 100.0%; Pred. No. 7.9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRFHLILVLOALPPATQGGKVVLGKKGDTVELTCTASOKSIOFHWKNSNOIK 60  
DB 1 MNRGVPRFHLILVLOALPPATQGGKVVLGKKGDTVELTCTASOKSIOFHWKNSNOIK 60  
QY ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGGSFLTKGPSKLNDRADSRSLMDQGNFPLIKKLTEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSDFHLLQGSGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
DB 121 LVFGLTANSDFHLLQGSGSLTLTLESPGSSPVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
QY 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208  
DB 181 TWTCTVLOKQKVEFKIDIVVLAFOKAS 208

RESULT 5  
AA38328  
ID AA38328 standard; protein; 394 AA.  
XX  
XX  
XX AAY8328;  
XX  
XX 14-JUL-2000 (first entry)  
XX  
XX T4 glycoprotein amino acid sequence.  
XX  
XX sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;  
KW AIDS; treatment; inhibit; cell to cell spread; infection; fusion.  
XX

OS Mammalia.  
 XX US5126433-A.  
 PN 30-JUN-1992.  
 XX 23-OCT-1987; 87US-00114244.  
 PD 21-AUG-1986; 86US-00898587.  
 PF (UYCO ) UNIV COLUMBIA NEW YORK.  
 XX Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;  
 PA WPI: 2000-348913/30.  
 XX N-PSDB; AAA10906.  
 DR Soluble T4 glycoprotein useful for prevention and treatment of acquired  
 XX immunodeficiency syndrome and for screening inhibitors of human  
 PT immunodeficiency viral binding.  
 PS Disclosure; Col 11-16; 64pp; English.  
 XX This sequence represents the full length amino acid sequence of  
 CC glycosylated sT4 glycoprotein. Human immunodeficiency virus (HIV) uses  
 CC sT4 as a target receptor on T cells. The invention relates to  
 CC glycosylated sT4 which functions by blocking the binding of HIV to T4  
 CC target cells, and can be used for the prophylaxis and treatment of AIDS  
 CC patients. Administration of sT4 effectively inhibits the cell to cell  
 CC spreading of HIV infection and also the fusion of HIV-infected T4 cells  
 CC and non-infected T4 cells. The administration of T4 alleviates several  
 CC symptoms associated with AIDS, and prevents the occurrence of new  
 CC pathological changes. The sT4 glycoprotein is useful for the prophylaxis  
 CC and treatment of patients with AIDS. It is also useful as a reagent to  
 CC identify natural, synthetic or recombinant molecules which act as  
 CC therapeutic agents or inhibitors of T4+ cell interactions and in  
 CC diagnostic assays for detection of T4 proteins or molecules  
 XX  
 SQ Sequence 394 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 394;  
 Best Local Similarity 100.0%; Pred. No. 7.9e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLVLTQALLPATQGGKVDVLTCTASQKKSIOFHMNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPATQGGKVDVLTCTASQKKSIOFHMNSNOIK 60  
 QY 61 ILNGQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILNGQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLTQNGKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCTVLTQNGKVEFKIDIVLAFQKAS 208  
 RESULT 6  
 AAB19509  
 ID AAB19509 standard; protein; 416 AA.  
 XX AAB19509;  
 XX 09-JAN-2001 (first entry)  
 XX CD4-IGM fusion protein CH4Mmu.  
 XX CD4; IGM; human; CD4Mmu; fusion protein; immunoglobulin; HIV; SIV; gp120;  
 KM therapy; diagnosis.

XX Homo sapiens.  
 OS Key  
 FH Protein  
 FT Protein  
 FT Protein  
 FT Protein  
 XX US6117656-A.  
 XX 12-SEP-2000.  
 PD 07-JUN-1995; 95US-00479353.  
 PF 22-JAN-1988; 88US-00147351.  
 PR 23-JAN-1989; 89US-00299596.  
 PR 09-JUN-1992; 92US-00896781.  
 PR 12-APR-1993; 93US-00057952.  
 PR 04-FEB-1994; 94US-00191708.  
 XX (GEHO ) GEN HOSPITAL CORP.  
 PA Seed B;  
 XX WPI: 2000-586558/55.  
 DR N-PSDB; AAA50662.  
 XX CD4-Immunoglobulin fusion proteins, useful for targeting gp120 of HIV or  
 PT SIV.  
 PS Example 1; Col 41-50; 39pp; English.  
 XX The present sequence is that of fusion protein CD4Mmu comprising the  
 CC extracellular portion of CD4, which binds to HIV gp120, linked at its C-  
 CC terminus to the human Igm heavy chain. To obtain the fusion protein, DNA  
 CC encoding CD4 was linked to Igm DNA at the Met2 site upstream of the CH1  
 CC region (see AAA50662). Fusion protein CD4Mmu and a nucleic acid encoding  
 CC it are claimed. Also claimed are a vector comprising the nucleic acid,  
 CC and a method of producing the fusion protein in secreted form using a  
 CC transformed host cell. The fusion protein may further comprise a  
 CC therapeutic agent, radiolabel or NMR imaging agent. The fusion protein  
 CC can be administered to an animal (including humans) for treatment of HIV  
 CC or SIV infection, and can also be used in assays for HIV or SIV, imaging  
 CC and tissue stains. Igm fusion proteins such as CD4Mmu provide complement-  
 CC mediated immunity  
 XX  
 SQ Sequence 416 AA;  
 Query Match 39.9%; Score 1064; DB 3; Length 416;  
 Best Local Similarity 100.0%; Pred. No. 8.3e-47;  
 Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPFRHLVLTQALLPATQGGKVDVLTCTASQKKSIOFHMNSNOIK 60  
 DB 1 MNRGVPFRHLVLTQALLPATQGGKVDVLTCTASQKKSIOFHMNSNOIK 60  
 QY 61 ILNGQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
 DB 61 ILNGQGSFLTKGSPSKLNDRADSRSLWDQGNFPLIINKLIEDSDTYICEVEDQKEVQL 120  
 QY 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTANSDTHLLQGSGLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWTCTVLTQNGKVEFKIDIVLAFQKAS 208  
 DB 181 TWTCTVLTQNGKVEFKIDIVLAFQKAS 208  
 RESULT 7  
 AAY51080  
 ID AAY51080 standard; protein; 436 AA.

```

XX AC AAY51080;
XX DT 23-MAR-2000 (first entry)
XX DE Human fusion protein CD4Mg.
XX DE Human fusion protein CD4Mg.
XX DE Fusion protein; human, CD4; Igm, immunoglobulin; gp120;
XX DE anti-human immunodeficiency virus, CD4Mg.
XX OS Homo sapiens.
XX OS Synthetic.
XX PN US6004781-A.
XX PD 21-DEC-1999.
XX PF 04-FEB-1994; 94US-00191708.
XX PR 22-JAN-1988; 88US-00147351.
XX PR 23-JAN-1989; 89US-00299596.
XX PR 09-JUN-1992; 92US-00896781.
XX PR 12-APR-1993; 93US-00057952.
XX PA (GENO ) GEN HOSPITAL CORP.
XX PI Seed B;
XX DR WPI; 2000-085792/07.
XX DR N-PSDB; AA244063.
XX PS Fusion protein useful for the treatment of human immunodeficiency virus.
XX PT Example 1; Col 41-50; 39pp; English.
XX CC This invention describes a novel nucleic acid (I) encoding a fusion
XX CC protein comprising a DNA sequence encoding amino acids 1-173 of CD4 (II)
XX CC and a DNA sequence encoding a human immunoglobulin (Ig) heavy or light
XX CC chain (III). The products of the invention have anti-human
XX CC immunodeficiency virus (HIV) activity and are capable of binding to
XX CC gp120. The fusion protein is useful for treating human immunodeficiency
XX CC virus (HIV) or simian immunodeficiency virus (SIV). This sequence
XX CC represents the fusion protein CD4Mg which is constructed from CD4 linked
XX CC to human Igm upstream of the CH1 region
XX SQ Sequence 436 AA;

Query Match 39.9%; Score 1064; DB 3; Length 436;
Best Local Similarity 100.0%; Pred. No. 8.7e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

XX DT 14-JUL-2000 (first entry)
XX DE T4 glycoprotein amino acid sequence.
XX DE T4 glycoprotein; human immunodeficiency virus; HIV; block binding;
XX DE sT4; glycoprotein; human immunodeficiency virus; HIV; block binding;
XX DE AIDS; treatment; inhibic; cell to cell spread; infection; fusion.
XX OS Mammalia.
XX PN US5126433-A.
XX PD 30-JUN-1992.
XX PF 23-OCT-1987; 87US-00114244.
XX PR 21-AUG-1986; 86US-00898587.
XX PA (UYCO ) UNIV COLUMBIA NEW YORK.
XX PI Maddon PJ, Chess L, Axel R, Weiss R, Littman DR, McDougal JS;
XX DR WPI; 2000-348913/30.
XX PT Soluble T4 glycoprotein useful for prevention and treatment of acquired
XX PT immunodeficiency syndrome and for screening inhibitors of human
XX PT immunodeficiency viral binding.
XX PS Example; Fig 6; 64pp; English.
XX CC This sequence represents the amino acid sequence of glycosylated sT4
XX CC glycoprotein. Human immunodeficiency virus (HIV) uses sT4 as a target
XX CC receptor on T cells. The invention relates to glycosylated sT4 which
XX CC functions by blocking the binding of HIV to T4 target cells, and can be
XX CC used for the prophylaxis and treatment of AIDS patients. Administration
XX CC of sT4 effectively inhibits the cell to cell spreading of HIV infection
XX CC and also the fusion of HIV-infected T4 cells and non-infected T4 cells.
XX CC The administration of T4 alleviates several symptoms associated with
XX CC AIDS, and prevents the occurrence of new pathological changes. The sT4
XX CC glycoprotein is useful for the prophylaxis and treatment of patients with
XX CC AIDS. It is also useful as a reagent to identify natural, synthetic or
XX CC recombinant molecules which act as therapeutic agents or inhibitors of
XX CC T4 cell interactions and in diagnostic assays for detection T4 proteins
XX CC or molecules
XX SQ Sequence 458 AA;

Query Match 39.9%; Score 1064; DB 3; Length 458;
Best Local Similarity 100.0%; Pred. No. 9e-47;
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

RESULT 8  
AAY88329  
ID AAY88329 standard; protein; 458 AA.  
XX  
AC AAY88329;

RESULT 9  
AAB81502  
ID AAB81502 standard; protein; 458 AA.  
XX  
AC AAB81502;

XX 18-JUN-2001 (first entry)  
DT Human CD4 protein.  
XX  
DE Human CD4 protein.  
XX  
KW Human; CD4; CD4 fusion protein; oligomerisation;  
KW receptor-ligand interaction inhibition; surface plasmon resonance; SPR;  
KW T cell receptor binding; MHC binding; carcinoma; autoimmune disease;  
KW multiple sclerosis; human immunodeficiency virus; HIV; diabetes;  
KW rheumatoid arthritis; immune disorder.  
XX  
OS Homo sapiens.  
XX  
FH Key  
FT Peptide 1..25  
FT Protein /label= signal\_peptide  
FT /label= 26..458  
FT /label= Human\_CD4  
XX  
PN WO200122084-A2.  
XX  
PD 29-MAR-2001.  
XX  
PF 18-SEP-2000; 2000WO-GB003579.  
XX  
PR 21-SEP-1999; 99GB-00022352.  
XX  
PA (AVID-) AVIDEX LTD.  
XX  
PI Jakobsen BK;  
XX  
XX WPI; 2001-273470/28.  
DR N-PSDB; AAF62582.  
XX  
PT Sequential screening of candidate compounds library for those which  
PT inhibit binding of low affinity receptor-ligand interaction having fast  
PT binding kinetics, using interfacial optical assay.  
XX  
PS Disclosure; Fig 13; 91pp; English.  
XX  
XX The present sequence is human CD4. Human CD4 extracellular domains 1 and  
XX 2 were used in the construction of CD4 oligomerisation fusion proteins.  
XX The fusion proteins contain an oligomerisation domain that enables the  
XX proteins to bind to one another to form oligomers. The oligomers may be  
XX used in an invention relating to a method for screening for compounds  
XX with the ability to inhibit a low affinity receptor-ligand interaction.  
XX The method uses an interfacial optical assay, such as surface plasmon  
XX resonance (SPR). The method is useful for screening candidate compounds  
XX for the ability to inhibit interaction between MHC/peptide complex and T  
XX cell receptor, and MHC/peptide complex and CD8 or CD4 co-receptor. The  
XX compounds identified by the above methods which interfere with T cell  
XX receptor binding to a particular HLA type molecule are useful as immune  
XX inhibitors for treating carcinomas, autoimmune diseases such as multiple  
XX sclerosis, human immunodeficiency virus (HIV) infection, rheumatoid  
XX arthritis, Hashimoto's disease, insulin dependent diabetes, Good  
XX pasture's syndrome, uveitis, psoriasis and graft rejection  
SQ Sequence 458 AA;  
Query Match 39.9%; Score 1064; DB 4; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVQLALPPATOGKKVYLGKGDVVELTCTASOKKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLLVQLALPPATOGKKVYLGKGDVVELTCTASOKKSIQFHKNSNQIK 60  
QY 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSPTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSPTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

DB 121 LVFGLTANSPTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
QY 181 TWICTVLQONQKVEFKIDIVLAFOKAS 208  
DB 181 TWICTVLQONQKVEFKIDIVLAFOKAS 208  
RESULT 10  
ID AAG79087  
AA AAG79087 standard; protein; 458 AA.  
XX  
AC AAG79087;  
XX  
DT 10-DEC-2001 (first entry)  
XX  
DE Amino acid sequence of human CD4 protein.  
XX  
KW Human; receptor; DC-SIGN; dendritic cell; T lymphocyte; HIV; gp120;  
KW C-type lectin; ICAM3; HIV entry; T cell; macrophage; HIV infection; CD4.  
XX  
OS Homo sapiens.  
XX  
PN WO200164752-A2.  
XX  
PD 07-SEP-2001.  
XX  
PF 28-FEB-2001; 2001WO-US006322.  
XX  
PR 02-MAR-2000; 2000US-00517605.  
XX  
PA (UYNY) UNIV NEW YORK STATE.  
XX  
PI (UYNI-) UNIV NIJMEGEN.  
XX  
PI Littman DR, Kwon D, Van Kooyk Y, Geijzenbeek T;  
XX  
XX WPI; 2001-602565/68.  
DR  
XX  
PT An antibody for the treatment or prevention of HIV-infection comprises a  
PT gp120 portion which binds to DC-SIGN or is exposed upon gp120 binding of  
PT DC-SIGN due to concomitant conformational change.  
XX  
XX Disclosure; Page 115-116; 131pp; English.  
PS  
XX The specification describes an antibody which is specific for an  
XX antigenic fragment of gp120. This antigenic fragment binds to DC-SIGN or  
XX is exposed upon gp120 binding of DC-SIGN due to concomitant  
XX conformational change. DC-SIGN is a receptor that is specifically  
XX expressed on dendritic cells and facilitates infection of T lymphocytes  
XX with HIV. DC-SIGN is identical to a HIV-1 gp120-binding C-type lectin. DC  
XX -SIGN binds ICAM-3 (which is expressed constitutively on T lymphocytes)  
XX with high affinity. The antibody of the invention inhibits the trans  
XX enhancement of HIV entry into a T cell or macrophage facilitated by  
XX dendritic cells. The antibody is useful to treat or prevent HIV  
XX infection. The present sequence represents a human CD4 protein  
SQ Sequence 458 AA;  
Query Match 39.9%; Score 1064; DB 4; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 MNRGVPFRHLLVQLALPPATOGKKVYLGKGDVVELTCTASOKKSIQFHKNSNQIK 60  
DB 1 MNRGVPFRHLLVQLALPPATOGKKVYLGKGDVVELTCTASOKKSIQFHKNSNQIK 60  
QY 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEVQL 120  
DB 61 ILNGSGFLTKGPKSLNDRADSRSLMDQGNPFLIIKNIKIEDSDTYICEVEDQKEVQL 120  
QY 121 LVFGLTANSPTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180  
DB 121 LVFGLTANSPTHLLOQGSILTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELQDSG 180

QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208  
 |||||  
 DB 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 11  
 ID ADD25609  
 ID ADD25609 standard; protein; 458 AA.  
 AC ADD25609;  
 XX  
 DT 15-JAN-2004 (first entry)  
 XX  
 DE Binding domain-immunoglobulin fusion protein-associated protein #82.  
 XX  
 KW Binding domain; immunoglobulin; fusion protein; cytostatic;  
 KW antitachytic; immunosuppressive; antidiabetic; antithyroid;  
 KW neuroprotective; hinge region; immunoglobulin heavy chain;  
 KW CH2 constant region; CH3 constant region; IgG1;  
 KW antibody dependent cell-mediated cytotoxicity; IgG1;  
 KW rheumatoid arthritis; myasthenia gravis; Grave's disease;  
 KW type 1 diabetes mellitus; multiple sclerosis; autoimmune disease.  
 XX  
 OS Unidentified.  
 PN US200318592-A1.  
 XX  
 PD 26-JUN-2003.  
 XX  
 PF 25-JUL-2002; 2002US-00207655.  
 XX  
 PR 17-JAN-2001; 2001US-0367358P.  
 PR 17-JAN-2002; 2002US-00053530.  
 PR 03-JUN-2002; 2002US-0385691P.  
 XX  
 PA (GENE-) GENE-CRAFT INC.  
 XX  
 PI Ledbetter JA, Hayden-Ledbetter MS, Thompson PA;  
 XX  
 DR WPI; 2003-801317/75.  
 XX  
 PT New binding domain-immunoglobulin fusion protein, useful for treating a  
 PT subject having or suspected of having a malignant condition or a B-cell  
 PT disorder, e.g., melanoma, Grave's disease or autoimmune disease.  
 XX  
 PS Disclosure; SEQ ID NO 170; 157bp; English.  
 XX  
 XX The invention relates to a binding domain-immunoglobulin fusion protein  
 CC comprising a binding domain polypeptide that is fused to an  
 CC immunoglobulin hinge region polypeptide, an immunoglobulin heavy chain  
 CC CH2 constant region polypeptide that is fused to the hinge region  
 CC polypeptide, and an immunoglobulin heavy chain CH3 constant region  
 CC polypeptide that is fused to the CH2 constant region polypeptide. The  
 CC hinge region polypeptide comprises: a wild-type human IgG1 immunoglobulin  
 CC hinge region polypeptide; a mutated human IgG1 immunoglobulin hinge  
 CC region polypeptide, derived from (a) having 3 or more cysteine residues;  
 CC where the mutated human IgG1 immunoglobulin hinge region polypeptide  
 CC contains 2 cysteine residues, where the first cysteine is not mutated; a  
 CC mutated human IgG1 immunoglobulin hinge region polypeptide, derived from  
 CC (a) having 3 or more cysteine residues, where the mutated human IgG1  
 CC immunoglobulin hinge region polypeptide contains no more than one  
 CC cysteine residue; and a mutated human IgG1 immunoglobulin hinge region  
 CC polypeptide, derived from (a) having 3 or more cysteine residues; where  
 CC the mutated human IgG1 immunoglobulin hinge region polypeptide contains  
 CC no cysteine residues. The binding domain-immunoglobulin fusion protein is  
 CC capable of at least one immunological activity comprising antibody  
 CC dependent cell-mediated cytotoxicity (ADCC) and complement fixation. The  
 CC binding domain polypeptide is capable of specifically binding to an  
 CC antigen. Also included are an isolated polynucleotide encoding the  
 CC binding domain-immunoglobulin fusion protein, a recombinant expression  
 CC construct comprising the polynucleotide (operably linked to a promoter),  
 CC a host cell transformed or transfected with a recombinant expression

CC construct, producing the binding domain-immunoglobulin fusion protein, a  
 CC pharmaceutical composition comprising the binding domain-immunoglobulin  
 CC fusion protein or polynucleotide and a carrier, and treating a subject  
 CC having or suspected of having a malignant condition or a B-cell disorder.  
 CC The binding domain-immunoglobulin fusion protein is useful for treating a  
 CC subject having or suspected of having a malignant condition or a B-cell  
 CC disorder, e.g., melanoma, carcinoma or sarcoma, rheumatoid arthritis,  
 CC myasthenia gravis, Grave's disease, type 1 diabetes mellitus, multiple  
 CC sclerosis or autoimmune disease. The present sequence is a binding domain  
 CC -immunoglobulin fusion protein-associated protein sequence. Note: The  
 CC sequence data for this patent formed part of the printed specification  
 CC and is also available in electronic format directly from USPTO at  
 CC [seqdata.uspto.gov/sequence.html?DocID=20030118592](http://seqdata.uspto.gov/sequence.html?DocID=20030118592). The authors have not  
 CC identified the sequences in the printed specification by their SEQ ID  
 CC number therefore none of the sequences can be explicitly identified.  
 XX  
 XX SQ Sequence 458 AA;  
 XX  
 QY Query Match 39.9%; Score 1064; DB 7; Length 458;  
 DB Best Local Similarity 100.0%; Pred. NO. 9e-47;  
 DB Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 MNRGVPRHLLVLTALLPATQKVVLGKGDVTELTCTASQKSIQFHMKNNSQIK 60  
 DB 1 MNRGVPRHLLVLTALLPATQKVVLGKGDVTELTCTASQKSIQFHMKNNSQIK 60  
 QY 61 ILNGQSFVTGPKSKLNRADSRSLMDQGNPLIINKLIEDSDTYICEVEDQKEEYOL 120  
 DB 61 ILNGQSFVTGPKSKLNRADSRSLMDQGNPLIINKLIEDSDTYICEVEDQKEEYOL 120  
 QY 121 LVFGLTNSDTHLLQGSGLTLTLESPPSSSVQCRSRGKNIQGGKTLVSQLELDQSG 180  
 DB 121 LVFGLTNSDTHLLQGSGLTLTLESPPSSSVQCRSRGKNIQGGKTLVSQLELDQSG 180  
 QY 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208  
 DB 181 TWCTCTVLQNKQKVEFKIDIVLAFQKAS 208

RESULT 12  
 ID ADE57489  
 ID ADE57489 standard; protein; 458 AA.  
 AC ADE57489;  
 DT 29-JAN-2004 (first entry)  
 XX  
 DE Human Protein P01730, SEQ ID NO 3351.  
 XX  
 KW Human; pain; neuronal tissue; gene therapy;  
 KW spinal segmental nerve injury; chronic constriction injury; CCI;  
 KW spared nerve injury; SNI; Chung.  
 OS Homo sapiens.  
 XX  
 XX WO2003016475-A2.  
 XX  
 PD 27-FEB-2003.  
 XX  
 PF 14-AUG-2002; 2002WO-US025765.  
 XX  
 PR 14-AUG-2001; 2001US-0312147P.  
 PR 01-NOV-2001; 2001US-0346382P.  
 PR 26-NOV-2001; 2001US-0333347P.  
 XX  
 PA (GENO) GEN HOSPITAL CORP.  
 PA (FAR) BAYER AG.  
 XX  
 PI Woolf C, D'urso D, Befort K, Costigan M;  
 XX  
 DR WPI; 2003-268312/26.  
 DR GENBANK; P01730.

PT New composition comprising two or more isolated polypeptides, useful for  
PT preparing a medicament for treating pain in an animal.

PS Claim 1, Page; 1017pp; English.

XX The invention discloses a composition comprising two or more isolated rat  
CC or human polynucleotides or a polynucleotide which represents a fragment,  
CC derivative or allelic variation of the nucleic acid sequence. Also  
CC claimed are a vector comprising the novel polynucleotide, a host cell  
CC comprising the vector, a method for identifying a nucleotide sequence  
CC which is differentially regulated in an animal subjected to pain and a  
CC kit to perform the method, an array, a method for identifying an agent  
CC that increases or decreases the expression of the polynucleotide sequence  
CC that is differentially expressed in neuronal tissue of a first animal  
CC subjected to pain, a method for identifying a compound which regulates  
CC the expression of a polynucleotide sequence which is differentially  
CC expressed in an animal subjected to pain, a method for identifying a  
CC compound that regulates the activity of one or more of the  
CC polynucleotides, a method for producing a pharmaceutical composition, a  
CC method for identifying a compound or small molecule that regulates the  
CC activity in an animal of one or more of the polypeptides given in the  
CC specification, a method for identifying a compound useful in treating  
CC pain and a pharmaceutical composition comprising the one or more  
CC polypeptides or their antibodies. The polynucleotide or the compound that  
CC modulates its activity is useful for preparing a medicament for treating  
CC pain (e.g., spinal segmental nerve injury (Chung), chronic constriction  
CC injury (CCI) and spared nerve injury (SNI)) in an animal (e.g., gene  
CC therapy). The sequence presented is a human protein (shown in Table 2 of  
CC the specification) which is differentially expressed during pain. Note:  
CC The sequence data for this patent did not form part of the printed  
CC specification, but was obtained in electronic form directly from WIPO at  
CC ftp.wipo.int/pub/published\_pct\_sequences.

XX Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 7; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRRHLLVQLALLPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRRHLLVQLALLPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNOIK 60  
QY 61 ILNGQGSFLTKGSPKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120  
DB 61 ILNGQGSFLTKGSPKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCVTVLQNGKVEFKIDIVLAFOKAS 208  
DB 181 TWTCVTVLQNGKVEFKIDIVLAFOKAS 208

RESULT 13

ADP12495

ID ADP12495 standard; protein; 458 AA.

XX ADP12495;

DT 12-AUG-2004 (first entry)

DE Protein encoded by mRNA of the invention #105.

XX transplamt rejection; immune system; rheumatoid arthritis; lupus;

KM inflammatory bowel disease; multiple sclerosis; HIV; AIDS.

OS Homo sapiens.

XX MO2004042346-A2.

XX

PD 21-MAY-2004.

XX 24-APR-2003; 2003MO-US012946.

XX 24-APR-2002; 2002US-00131831.

PR 20-DEC-2002; 2002US-00325899.

XX (EXPR-) EXPRESSION DIAGNOSTICS INC.

PI Wohlgenuth J, Fry K, Woodward R, Ly N, Prentice J, Morris M;

PI Rosenberg S;

DR WPI; 2004-400724/37.

XX Diagnosing or monitoring transplant rejection, e.g. heart, kidney, liver,  
PT pancreas, pancreatic islet, lung, bone marrow or stem cell transplant  
PT rejection, in an individual, comprises detecting the expression level of  
PT the genes.

XX Claim 65; SEQ ID NO 2504; 1762pp; English.

CC The present invention relates to diagnosing or monitoring transplant  
CC rejection, e.g. cardiac or kidney transplant rejection, in an individual  
CC comprising detecting the expression level of one or more genes. The  
CC methods, system and kits are useful in diagnosing or monitoring  
CC transplant rejection, e.g. heart, kidney, liver, pancreas, pancreatic  
CC islet, lung, bone marrow or stem cell transplant rejection,  
CC xenotransplant rejection or mechanical organ replacement rejection, in an  
CC individual. The method is also useful in assessing the immune status of  
CC an individual. The methods are also useful in diagnosing and monitoring  
CC diseases that involve the immune system, e.g. rheumatoid arthritis,  
CC lupus, inflammatory bowel diseases, multiple sclerosis, HIV/AIDS or  
CC viral, bacterial or fungal infection. The present sequence represents a  
CC protein that is encoded by the mRNA of the invention.

XX Sequence 458 AA:

Query Match 39.9%; Score 1064; DB 8; Length 458;  
Best Local Similarity 100.0%; Pred. No. 9e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNRGVPRRHLLVQLALLPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNOIK 60  
DB 1 MNRGVPRRHLLVQLALLPAATQGGKVVGGKGDVVELTCTASOKKSIOFHMKNSNOIK 60  
QY 61 ILNGQGSFLTKGSPKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120  
DB 61 ILNGQGSFLTKGSPKLNDRADSRSLMDQGNPFLIKNLKIEDSDTYICEVEDQKEVOL 120  
QY 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
DB 121 LVFGLTANSDTHLLQGGSLTLTLESPGSSPSVQCRSPRGKNIQGGKTLVSQLELDGSG 180  
QY 181 TWTCVTVLQNGKVEFKIDIVLAFOKAS 208  
DB 181 TWTCVTVLQNGKVEFKIDIVLAFOKAS 208

RESULT 14

ADA44807

ID ADA44807 standard; protein; 473 AA.

XX ADA44807;

DT 04-DEC-2003 (first entry)

DE CD4/TCR CD3epsilon chain chimeric protein CD4epsilon10n15, SEQ ID NO:2.

XX HIV-1 infection; human immunodeficiency virus-1; CD4+ cell; chimeric CD4;

KM endoplasmic reticulum; ER retention; envelope protein gp160;

KM T cell receptor CD3epsilon chain; C-terminal domain; CD4epsilon10n15;

KM gene therapy; human; receptor.

XX

OS Chimeric.  
OS Homo sapiens.  
FH Key Location/Qualifiers  
FT Protein 1..458  
FT /label= CD4  
FT Region 459..473  
FT /note= "Part of the C-terminal domain of the T cell  
FT receptor CD3epsilon chain"  
XX  
XX WO2003076468-A1.  
XX  
XX 18-SEP-2003.  
XX  
XX 14-MAR-2003; 2003WO-ES000120.  
XX  
XX 14-MAR-2002; 2002ES-00000616.  
XX  
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.  
XX  
XX Alarcon Sanchez BJ, San Jose Martinez ME, Zaldivar Notario I;  
XX Gomez Buendia M;  
XX  
XX WPI; 2003-779059/73.  
XX  
XX N-PSDB; ADA44806.  
XX  
XX  
XX Composition for treating or preventing human immune deficiency virus,  
XX comprises CD4 chimeric protein having a protective effect in trans, or  
XX related nucleic acid.  
XX  
XX  
XX Claim 5, Page 33-35; 43pp; Spanish.  
XX  
XX The invention relates to a composition for the treatment or prevention of  
XX human immunodeficiency virus-1 (HIV-1) infection. The composition  
XX comprises CD4+ cells that have been transduced with a vector that encodes  
XX a chimeric CD4 molecule which is capable of being retained in the  
XX endoplasmic reticulum (ER). The invention also encompasses the use of a  
XX soluble protein factor produced by CD4+ cells that have been transduced  
XX with a vector encoding a chimeric CD4 protein; and the use of an  
XX expression system encoding a chimeric CD4 protein. The ER-localised  
XX chimeric CD4 molecule binds to the HIV-1 envelope protein gp160,  
XX resulting in HIV-1 retention in the ER and thereby preventing viral  
XX replication. In a specific embodiment, the chimeric CD4 molecule  
XX comprises CD4 fused to 15 amino acids of the C-terminal domain of the T  
XX cell receptor CD3epsilon chain; this chimeric CD4 molecule is designated  
XX CD4epsilon10n5 (ADA44807). A known chimeric CD4 of similar structure but  
XX containing only 10 amino acids from CD3epsilon can also be used.  
XX Compositions of the invention have an in trans effect on the replication  
XX of HIV-1, and may be used to treat and prevent HIV-1 infection. The  
XX present sequence represents the chimeric CD4 molecule CD4epsilon10n5, which  
XX is specifically claimed for use in compositions of the invention.  
XX  
XX  
SQ Sequence 473 AA;  
Query Match 39.9%; Score 1064; DB 7; Length 473;  
Best Local Similarity 100.0%; Pred. No. 9.3e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 15  
AA59170  
ID AA59170 standard; protein; 474 AA.  
XX  
XX  
XX AA59170;  
AC  
XX  
XX 14-MAR-2000 (first entry)  
DT  
XX  
XX CD4-Ig fusion protein CD4mu.  
DE  
XX  
XX HIV; extracellular; CD4; gp120; immunoglobulin; Ig; fusion protein;  
XX secreted protein; SIV infection; medicament.  
XX  
XX Synthetic.  
OS Homo sapiens.  
XX  
XX CA1340741-C.  
XX  
XX  
XX 14-SEP-1999.  
XX  
XX  
XX 20-JAN-1989; 89CA-00588749.  
XX  
XX 20-JAN-1989; 89CA-00588749.  
XX  
XX (GENO ) GEN HOSPITAL CORP.  
XX  
XX Seed B;  
XX  
XX  
XX WPI; 2000-061015/06.  
XX  
XX N-PSDB; AA248203.  
XX  
XX  
XX New fusion gene encoding immunoglobulin-CD4 fusion proteins, useful in  
XX the treatment of HIV or simian immunodeficiency virus infections.  
XX  
XX  
XX Example 1; Page 47-53; 89pp; English.  
XX  
XX The invention provides a fusion gene encoding a fusion protein that  
XX comprises an extracellular CD4 DNA sequence or its fragment which binds  
XX to HIV gp120 when fused to an immunoglobulin (Ig) chain and the DNA  
XX sequence of an Ig heavy or light chain, where the DNA sequence encoding  
XX the variable region has been replaced with the DNA sequence which encodes  
XX extracellular CD4 or its gp120 binding fragment. The fusion protein is  
XX capable of being secreted. The fusion proteins are useful for treating  
XX HIV or SIV infections in animals, preferably humans. They are also useful  
XX for producing medicaments which can be used for treating HIV or SIV  
XX infections in humans. The present sequence represents the fusion protein  
XX CD4mu where the CD4 is linked to human IgG1 at the Met2 site upstream of  
XX the CH1 region  
XX  
XX  
SQ Sequence 474 AA;  
Query Match 39.9%; Score 1064; DB 3; Length 474;  
Best Local Similarity 100.0%; Pred. No. 9.3e-47;  
Matches 208; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Sun Nov 28 09:38:24 2004

us-09-936-702-3.rag

Page 10

Search completed: November 21, 2004, 13:37:17  
Job time : 265.343 secs

---